

日本国特許庁
JAPAN PATENT OFFICE

02.08.2004

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

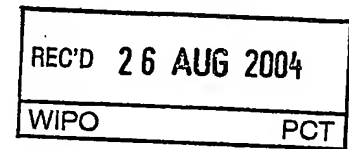
This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application: 2003年 7月29日

出願番号
Application Number: 特願2003-282122

[ST. 10/C]: [JP2003-282122]

出願人
Applicant(s): 日本電気株式会社

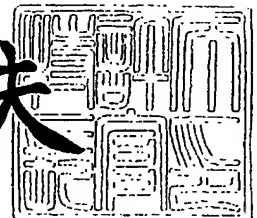


PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

2004年 5月25日

特許庁長官
Commissioner,
Japan Patent Office

今井康夫



【書類名】 特許願
【整理番号】 34103785
【提出日】 平成15年 7月29日
【あて先】 特許庁長官
【国際特許分類】 C12Q 1/68
【発明者】
 【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内
 【氏名】 喜友名 朝春
【発明者】
 【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内
 【氏名】 岡島 健治
【発明者】
 【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内
 【氏名】 鳥居 宏在
【発明者】
 【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内
 【氏名】 上條 憲一
【特許出願人】
 【識別番号】 000004237
 【氏名又は名称】 日本電気株式会社
【代理人】
 【識別番号】 100110928
 【弁理士】
 【氏名又は名称】 速水 進治
 【電話番号】 03-5784-4637
【手数料の表示】
 【予納台帳番号】 138392
 【納付金額】 21,000円
【提出物件の目録】
 【物件名】 特許請求の範囲 1
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1
 【包括委任状番号】 0110433

【書類名】 特許請求の範囲**【請求項 1】**

染色体テリトリを含む細胞を撮影した画像情報に基づいて、細胞の状態を評価する方法であって、

前記画像から、前記染色体テリトリを抽出するステップと、

前記染色体テリトリの配置状態を標準化した後、当該配置状態を定量化するステップと、

定量化された前記染色体テリトリの配置状態に基づき、前記細胞の状態を評価するステップと、

を含むことを特徴とする評価方法。

【請求項 2】

請求項 1 に記載の評価方法において、

前記染色体テリトリを抽出するステップにおいて、前記細胞の核を抽出するとともに当該核内に含まれる前記染色体テリトリを抽出し、

前記配置状態を定量化するステップは、前記核の形状を標準化するステップと、標準化前の前記核および標準化後の前記核の形状変化に基づいて、前記染色体テリトリの座標を変換するステップと、を含むことを特徴とする評価方法。

【請求項 3】

請求項 1 または 2 に記載の評価方法において、

複数の細胞の前記配置状態を定量化した結果を統計処理するステップをさらに含み、

前記細胞の状態を評価するステップにおいて、一の細胞の前記配置状態を定量化した結果と前記統計処理の結果とに基づき、前記一の細胞の状態を評価することを特徴とする評価方法。

【請求項 4】

請求項 3 に記載の評価方法において、

前記統計処理するステップは、複数の細胞の前記配置状態を定量化した結果に基づき、前記配置状態の定量化結果の標準値を算出するステップを含み、

前記細胞の状態を評価するステップにおいて、一の細胞の前記配置状態を定量化した結果と前記標準値とに基づき、前記一の細胞の状態を評価することを特徴とする評価方法。

【請求項 5】

染色体テリトリを含む細胞を撮影した画像情報に基づいて、細胞の状態を評価する方法であって、

前記画像から、複数の染色体テリトリをそれぞれ抽出するステップと、

前記複数の染色体テリトリの配置状態を定量化するステップと、

複数の細胞の前記配置状態を定量化した結果を統計処理するステップと、

一の細胞の前記配置状態を定量化した結果と、前記統計処理の結果とに基づき、前記一の細胞の状態を評価するステップと、

を含むことを特徴とする評価方法。

【請求項 6】

請求項 5 に記載の評価方法において、

前記配置状態は、前記複数の染色体テリトリ間の距離、または前記複数の染色体テリトリの配置方向の少なくともいずれか 1 つを含むことを特徴とする評価方法。

【請求項 7】

請求項 5 または 6 に記載の評価方法において、

前記定量化するステップは、前記複数の染色体テリトリ間の距離を算出するステップを含むことを特徴とする評価方法。

【請求項 8】

請求項 7 に記載の評価方法において、

前記距離を算出するステップは、

前記複数の染色体テリトリそれぞれの重心を算出するステップと、

前記重心間の距離を算出するステップと、
を含むことを特徴とする評価方法。

【請求項 9】

請求項 5 乃至 8 いずれかに記載の評価方法において、
前記定量化するステップは、前記複数の染色体テリトリの配置方向を定量化するステップを含むことを特徴とする評価方法。

【請求項 10】

請求項 9 に記載の評価方法において、
前記配置方向を定量化するステップは、
前記複数の染色体テリトリの主軸をそれぞれ検出するステップと、
前記主軸の角度を算出するステップと、
を含むことを特徴とする評価方法。

【請求項 11】

染色体テリトリを含む細胞を撮影した画像情報に基づいて、細胞の状態を評価する方法であって、
前記画像から、前記細胞の核および前記染色体テリトリをそれぞれ抽出するステップと

、
前記核および前記染色体テリトリの配置状態を定量化するステップと、
複数の細胞の前記配置状態を定量化した結果を統計処理するステップと、
一の細胞の前記配置状態を定量化した結果と、前記統計処理の結果とに基づき、前記一の細胞の状態を評価するステップと、
を含むことを特徴とする評価方法。

【請求項 12】

請求項 11 に記載の評価方法において、
前記定量化するステップは、前記核内の基準点と前記染色体テリトリとの間の距離を算出するステップを含むことを特徴とする評価方法。

【請求項 13】

請求項 5 乃至 12 いずれかに記載の評価方法において、
前記染色体テリトリの配置状態を定量化するステップは、前記染色体テリトリの配置状態を標準化するステップをさらに含み、前記染色体テリトリの配置状態を標準化した後に、当該配置状態を定量化することを特徴とする評価方法。

【請求項 14】

請求項 13 に記載の評価方法において、
前記統計処理するステップは、複数の細胞の前記配置状態を定量化した結果に基づき、前記配置状態の定量化結果の標準値を算出するステップを含み、
前記細胞の状態を評価するステップにおいて、一の細胞の前記配置状態を定量化した結果と、前記標準値とに基づき、前記一の細胞の状態を評価することを特徴とする評価方法。

【請求項 15】

請求項 1 乃至 14 いずれかに記載の評価方法において、
前記配置状態は、前記細胞の核内における前記染色体テリトリの相対位置、前記染色体テリトリの主軸の方向、および前記染色体テリトリの広がり少なくともいずれか 1 つを含むことを特徴とする評価方法。

【請求項 16】

請求項 1 乃至 15 いずれかに記載の評価方法において、
前記画像は、それぞれ属性値を有する複数のピクセルにより構成され、
前記染色体テリトリを抽出するステップは、
前記画像を複数のクラスに分類するステップと、
前記クラスのいずれかを前記染色体テリトリを表す領域として前記画像から抽出するステップと、を含み、

前記分類するステップは、
前記クラスそれぞれの属性値を示す属性パラメータおよび前記クラスそれぞれの混合比率の初期値を設定するステップと、
前記属性パラメータおよび前記混合比率に基づき、前記複数のピクセルのそれぞれが前記クラスそれぞれに分類されるクラス帰属確率を算出するステップと、
前記属性パラメータおよび前記混合比率により規定される混合確率分布、および前記帰属確率に基づき、推定の良さを示す評価関数を算出し、当該評価関数が所定の条件を満たすか否かを判定するステップと、
前記評価関数が前記所定の条件を満たさない場合、前記属性パラメータおよび前記混合比率を更新するステップと、
前記評価関数が前記所定の条件を満たすときの前記属性パラメータおよび前記混合比率に基づき、前記複数のピクセルを前記複数のクラスのいずれかに分類するステップと、
を含むことを特徴とする評価方法。

【請求項 1 7】

請求項 1 6 に記載の評価方法において、
前記クラス帰属確率を算出するステップは、
前記複数のピクセルをその属性値に応じて複数の部分空間に分解するステップと、
前記部分空間に含まれるピクセルの割合を示す粗視化経験確率分布を算出するステップと、
前記属性パラメータおよび前記混合比率に基づき、前記部分空間のそれぞれが前記クラスそれぞれに分類される粗視化クラス帰属確率を算出するステップと、を含み、
前記評価関数を算出するステップは、前記属性パラメータおよび前記混合比率により規定される混合確率分布を前記部分空間内で平均化して粗視化混合確率分布を算出するステップを含み、前記粗視化経験確率分布、前記粗視化混合確率分布、および前記粗視化クラス帰属確率に基づき、前記評価関数を算出することを特徴とする評価方法。

【請求項 1 8】

染色体テリトリを含む細胞を撮影した画像情報に基づいて、細胞の状態を評価するシステムであって、
前記画像から、前記染色体テリトリを抽出する抽出処理部と、
抽出した前記染色体テリトリの配置状態を定量化する定量化処理部と、
前記配置状態を定量化した結果を記憶する記憶部と、
複数の細胞の前記配置状態を定量化した結果を統計処理する統計処理部と、
一の細胞の前記配置状態を定量化した結果と、前記統計処理の結果とに基づき、前記一の細胞の状態を評価する評価部と、
を含むことを特徴とする評価システム。

【請求項 1 9】

請求項 1 8 に記載の評価システムにおいて、
前記定量化処理部は、前記画像から抽出した前記染色体テリトリの配置状態を標準化し、前記染色体テリトリの配置状態を標準化した後に、当該配置状態を定量化することを特徴とする評価システム。

【請求項 2 0】

請求項 1 9 に記載の評価システムにおいて、
前記抽出処理部は、前記細胞の核を抽出するとともに当該核内に含まれる前記染色体テリトリを抽出し、
前記定量化処理部は、前記核の形状を標準化し、標準化前の前記核および標準化後の前記核の形状変化に基づいて、前記染色体テリトリの座標を変換することを特徴とする評価システム。

【請求項 2 1】

コンピュータに、染色体テリトリを含む細胞を撮影した画像情報に基づいて、細胞の状態を評価させるためのプログラムであって、

コンピュータを、
前記画像から、前記染色体テリトリを抽出する手段、
前記染色体テリトリの配置状態を標準化した後、当該配置状態を定量化する手段、
定量化された前記染色体テリトリの配置状態に基づき、前記細胞の状態を評価する手段
、
として機能させることを特徴とするプログラム。

【請求項 22】

コンピュータに、染色体テリトリを含む細胞を撮影した画像情報に基づいて、細胞の状態を評価させるためのプログラムであって、
コンピュータを、
前記画像から、前記染色体テリトリを抽出する抽出処理手段、
抽出した前記染色体テリトリの配置状態を定量化する定量化処理手段、
前記配置状態を定量化した結果を記憶する記憶手段、
複数の細胞の前記配置状態を定量化した結果を統計処理する統計処理手段、
一の細胞の前記配置状態を定量化した結果と、前記統計処理の結果とに基づき、前記一の細胞の状態を評価する評価手段、
として機能させることを特徴とするプログラム。

【書類名】明細書

【発明の名称】染色体状態の評価方法および評価システム

【技術分野】

【0001】

本発明は、細胞周期の間期における染色体テリトリの配置状態に基づき、細胞および細胞に含まれる染色体の状態を評価する方法およびシステムに関する。

【背景技術】

【0002】

従来、染色体の形態は細胞周期の分裂期における凝縮した状態でしか観測することができなかった。しかし、近年の可視化技術（3D-FISH、共焦点レーザー顕微鏡）の発達により、細胞周期の間期における染色体の配置や形態を観測することが可能になってきた。このような可視化技術により、間期における細胞の核内において、個々の染色体が互いに混ざり合うことのない領域に分かれて存在することが明らかになってきた（非特許文献1、非特許文献2）。以下、個々の染色体が核内において占有する領域を染色体テリトリ(chromosome territory)という。

【非特許文献1】T. Cremer, C. Cremer, CHROMOSOME TERRITORIES NUCLEAR ARCHITECTURE AND VENE REGULATION IN MAMMALIAN CELLS, NATURE REVIEWS | GENETICS, vol. 2, pp. 292-301, 2001

【非特許文献2】田辺秀之、「染色体テリトリ：間期核における染色体の核内配置と核高次構造に関する最近の研究」、Environ. Mutagen Res., 25, 11-22頁、2003

【特許文献1】特開2001-92980号公報

【発明の開示】

【発明が解決しようとする課題】

【0003】

ところで、正常細胞から腫瘍細胞への変化、および良性腫瘍細胞から悪性腫瘍細胞への変化は、遺伝子の変異や遺伝子の転座によっておこると考えられている。非特許文献2には、空間的に近接して配置されている染色体テリトリどうしは、互いに遠方に配置されているものに比べ、転座を起こす確率が高いという考え方が記載されている。

【0004】

以上のことから、染色体テリトリどうしの空間配置を検討することにより、細胞の変化を評価することが期待される。しかし、従来、染色体テリトリどうしの空間配置を定量的に解析し、細胞の変化を評価する技術は提供されていない。

【0005】

本出願の発明者らは、細胞周期の間期における染色体テリトリの配置状態を調べることで、細胞の状態を評価し、評価結果に基づき、癌等の疾患の診断の指標とすることを目的とし、以下の発明を考案した。

【課題を解決するための手段】

【0006】

本発明によれば、染色体テリトリを含む細胞を撮影した画像情報に基づいて、細胞の状態を評価する方法であって、画像から、染色体テリトリを抽出するステップと、染色体テリトリの配置状態を標準化した後、当該配置状態を定量化するステップと、定量化された染色体テリトリの配置状態に基づき、細胞の状態を評価するステップと、を含むことを特徴とする評価方法が提供される。

【0007】

ここで、配置状態は、細胞の核内における染色体テリトリの相対位置、染色体テリトリの主軸の方向、および染色体テリトリの広がり少なくともいずれか1つを含むことができる。また、標準化とは、複数の細胞における染色体領域の形状を統一した指標で定量化

するための処理である。このような処理を行うことにより、複数の細胞について、統一した指標で定量化を行うことができる。

【0008】

ここで、細胞の状態を評価するとは、細胞に含まれる染色体状態を評価すること、ならびに細胞に生じた変化、細胞の形態異常の予期、染色体転座、および細胞に生じた変化に基づく癌等の疾患に罹患している可能性を評価することをいう。本発明の評価方法によれば、染色体テリトリの形状を定量化することができるので、定量化した結果に基づき、細胞の形態異常や疾患の有無を早期に検知することが期待できる。

【0009】

細胞を撮影した画像は、細胞の核を核断面で画像化した複数の画像とすることができ、染色体テリトリを抽出するステップは、複数の画像から染色体テリトリをそれぞれ抽出し、これらを合成して染色体テリトリ像を3次元で抽出することができる。

【0010】

本発明の評価方法において、染色体テリトリを抽出するステップにおいて、細胞の核を抽出するとともに当該核内に含まれる染色体テリトリを抽出することができ、配置状態を定量化するステップは、核の形状を標準化するステップと、標準化前の核および標準化後の核の形状変化に基づいて、染色体テリトリの座標を変換するステップと、を含むことができる。ここで、核の形状変化とは、かたがちが変化すること、および大きさが変化することのいずれも含んでよい。染色体テリトリの座標を変換するステップは、標準化前の核と染色体テリトリとの相対関係、および標準化後の核と染色体テリトリとの相対関係に基づいて座標変換を行うことができる。

【0011】

本発明の評価方法は、複数の細胞の配置状態を定量化した結果を統計処理するステップをさらに含むことができ、細胞の状態を評価するステップにおいて、一の細胞の配置状態を定量化した結果と統計処理の結果とに基づき、一の細胞の状態を評価することができる。

【0012】

本発明の評価方法において、総計処理するステップは、複数の細胞の配置状態を定量化した結果に基づき、配置状態の定量化結果の標準値を算出するステップを含むことができ、細胞の状態を評価するステップにおいて、一の細胞の配置状態を定量化した結果と標準値とに基づき、一の細胞の状態を評価することができる。

【0013】

本発明の評価方法において、標準値を算出するステップは、複数の細胞の染色体の配置を定量化した結果の平均値を算出するステップを含むことができる。

【0014】

本発明によれば、染色体テリトリを含む細胞を撮影した画像情報に基づいて、細胞の状態を評価する方法であって、画像から、複数の染色体テリトリをそれぞれ抽出するステップと、複数の染色体テリトリの配置状態を定量化するステップと、複数の細胞の配置状態を定量化した結果を統計処理するステップと、一の細胞の配置状態を定量化した結果と、統計処理の結果とに基づき、一の細胞の状態を評価するステップと、を含むことを特徴とする評価方法が提供される。

【0015】

このように一の細胞の染色体テリトリの配置状態を定量化した結果を統計処理の結果と比較することにより、その一の細胞の異常度を評価することができる。また、本発明の評価方法によれば、染色体テリトリの形状を定量化することができるので、定量化した結果に基づき、細胞の形態異常や疾患の有無を早期に検知することが期待できる。

【0016】

本発明の評価方法において、配置状態は、複数の染色体テリトリ間の距離、または複数の染色体テリトリの配置方向の少なくともいずれか1つを含むことができる。

【0017】

本発明の評価方法において、定量化するステップは、複数の染色体テリトリ間の距離を算出するステップを含むことができる。

【0018】

本発明の評価方法において、距離を算出するステップは、複数の染色体テリトリそれぞれの重心を算出するステップと、重心間の距離を算出するステップと、を含むことができる。

【0019】

また、本発明の評価方法において、距離を算出するステップは、以下のようにすることもできる。

【0020】

本発明の評価方法において、距離を算出するステップは、マハラノビスの距離を用いて染色体テリトリ間の距離を算出するステップを含むことができる。

【0021】

本発明の評価方法において、距離を算出するステップは、染色体テリトリの境界間の距離を算出するステップを含むことができる。ここで、染色体テリトリの境界間の距離としては、境界間の最短距離を用いることができる。

【0022】

本発明の評価方法において、距離を算出するステップは、複数の染色体テリトリの形状をそれぞれ正規分布で近似するステップと、染色体テリトリそれぞれの正規分布の平均値を算出するステップと、平均値間の距離を算出するステップと、を含むことができる。

【0023】

本発明の評価方法において、距離を算出するステップは、複数の染色体テリトリそれぞれの骨格線を検出するステップと、骨格線の重心を算出するステップと、重心間の距離を算出するステップと、を含むことができる。

【0024】

本発明の評価方法において、前記定量化するステップは、前記複数の染色体テリトリの配置方向を定量化するステップを含むことができる。

【0025】

本発明の評価方法において、前記配置方向を定量化するステップは、前記複数の染色体テリトリの主軸をそれぞれ検出するステップと、前記主軸の角度を算出するステップと、を含むことができる。

【0026】

本発明の評価方法において、主軸を検出するステップは、複数の染色体テリトリの形状を正規分布で近似するステップと、近似した染色体テリトリの形状に基づき、染色体テリトリそれぞれの正規分散の平均値および共分散行列を算出するステップと、共分散行列を固有値分解するステップと、を含むことができる。

【0027】

本発明によれば、染色体テリトリを含む細胞を撮影した画像情報に基づいて、細胞の状態を評価する方法であって、画像から、細胞の核および染色体テリトリをそれぞれ抽出するステップと、核および染色体テリトリの配置状態を定量化するステップと、複数の細胞の配置状態を定量化した結果を統計処理するステップと、一の細胞の配置状態を定量化した結果と、統計処理の結果とに基づき、一の細胞の状態を評価するステップと、を含むことを特徴とする評価方法が提供される。

【0028】

本発明の評価方法において、定量化するステップは、核内の基準点と染色体テリトリとの間の距離を算出するステップを含むことができる。ここで、核内の基準点とは、たとえば、核壁、または重心とすることができる。

【0029】

本発明の評価方法は、染色体テリトリの配置状態を定量化するステップは、染色体テリトリの配置状態を標準化するステップをさらに含み、染色体テリトリの配置状態を標準化

した後に、当該配置状態を定量化することができる。

【0030】

本発明の評価方法において、統計処理するステップは、複数の細胞の配置状態を定量化した結果に基づき、配置状態の定量化結果の標準値を算出するステップを含むことができ、細胞の状態を評価するステップにおいて、一の細胞の配置状態を定量化した結果と、標準値とに基づき、一の細胞の状態を評価することができる。

【0031】

本発明の評価方法において、標準値を算出するステップは、複数の細胞の配置状態を定量化した結果の平均値を算出するステップを含むことができる。

【0032】

本発明によれば、染色体テリトリを含む細胞を撮影した画像情報に基づいて、細胞の状態を評価する方法であって、画像から、細胞の核および複数の染色体テリトリをそれぞれ抽出するステップと、核の形状を標準化するステップと、染色体テリトリの座標を、標準化前の核および標準化後の核の形状変化に応じて定められた一定の基準に従って変換するステップと、変換された染色体テリトリの重心をそれぞれ算出するステップと、複数の染色体テリトリの重心間の距離を算出するステップと、複数の細胞について、得られた重心間の距離を統計処理するステップと、一の細胞の重心間の距離と、統計処理の結果とに基づき、一の細胞の状態を評価するステップと、を含むことを特徴とする評価方法が提供される。

【0033】

本発明の評価方法において、細胞の状態を評価するステップは、t検定を用いて一の細胞の配置状態を定量化した結果と統計処理の結果との差異を定量化するステップを含むことができる。

【0034】

本発明の評価方法において、細胞の状態を評価するステップは、カイ2乗検定を用いて一の細胞の配置状態を定量化した結果と統計処理の結果との差異を定量化するステップを含むことができる。

【0035】

本発明によれば、染色体テリトリを含む細胞を撮影した画像情報に基づいて、細胞の状態を評価するシステムであって、画像から、染色体テリトリを抽出する抽出処理部と、抽出した染色体テリトリの配置状態を定量化する定量化処理部と、配置状態を定量化した結果を記憶する記憶部と、複数の細胞の配置状態を定量化した結果を統計処理する統計処理部と、一の細胞の配置状態を定量化した結果と、統計処理の結果とに基づき、一の細胞の状態を評価する評価部と、を含むことを特徴とする評価システムが提供される。

【0036】

本発明の評価システムにおいて、定量化処理部は、画像から抽出した染色体テリトリの配置状態を標準化し、染色体テリトリの配置状態を標準化した後に、当該配置状態を定量化することができる。

【0037】

本発明の評価システムにおいて、抽出処理部は、細胞の核を抽出するとともに当該核内に含まれる染色体テリトリを抽出することができ、定量化処理部は、核の形状を標準化し、標準化前の核および標準化後の核の形状変化に基づいて、染色体テリトリの座標を変換することができる。

【0038】

本発明によれば、コンピュータに、染色体テリトリを含む細胞を撮影した画像情報に基づいて、細胞の状態を評価させるためのプログラムであって、コンピュータを、画像から、染色体テリトリを抽出する手段、染色体テリトリの配置状態を標準化した後、当該配置状態を定量化する手段、定量化された染色体テリトリの配置状態に基づき、細胞の状態を評価する手段、として機能させることを特徴とするプログラムが提供される。

【0039】

本発明によれば、コンピュータに、染色体テリトリを含む細胞を撮影した画像情報に基づいて、細胞の状態を評価させるためのプログラムであって、コンピュータを、画像から、染色体テリトリを抽出する抽出処理手段、抽出した染色体テリトリの配置状態を定量化する定量化処理手段、配置状態を定量化した結果を記憶する記憶手段、複数の細胞の配置状態を定量化した結果を統計処理する統計処理手段、一の細胞の配置状態を定量化した結果と、統計処理の結果とに基づき、一の細胞の状態を評価する評価手段、として機能させることを特徴とするプログラムが提供される。

【0040】

本発明の評価方法において、画像は、それぞれ属性値を有する複数のピクセルにより構成され、染色体テリトリを抽出するステップは、画像を複数のクラスに分類するステップと、クラスのいずれかを染色体テリトリを表す領域として画像から抽出するステップと、を含むことができる。ここで、分類するステップは、クラスそれぞれの属性値を示す属性パラメータおよびクラスそれぞれの混合比率の初期値を設定するステップと、属性パラメータおよび混合比率に基づき、複数のピクセルのそれぞれがクラスそれぞれに分類されるクラス帰属確率を算出するステップと、属性パラメータおよび混合比率により規定される混合確率分布、および帰属確率に基づき、推定の良さを示す評価関数を算出し、当該評価関数が所定の条件を満たすか否かを判定するステップと、評価関数が所定の条件を満たさない場合、属性パラメータおよび混合比率を更新するステップと、評価関数が所定の条件を満たすときの属性パラメータおよび混合比率に基づき、複数のピクセルを複数のクラスのいずれかに分類するステップと、を含むことを特徴とする。

【0041】

この方法を用いることにより、画像から染色体テリトリを精度良く抽出することができる。染色体テリトリを精度良く抽出することにより、染色体テリトリの定量化結果のばらつきを低減することができ、細胞の状態を良好に評価することができる。

【0042】

本発明の評価方法において、クラス帰属確率を算出するステップは、複数のピクセルをその属性値に応じて複数の部分空間に分解するステップと、部分空間に含まれるピクセルの割合を示す粗視化経験確率分布を算出するステップと、属性パラメータおよび混合比率に基づき、部分空間のそれぞれがクラスそれぞれに分類される粗視化クラス帰属確率を算出するステップと、を含むことができ、評価関数を算出するステップは、属性パラメータおよび混合比率により規定される混合確率分布を部分空間内で平均化して粗視化混合確率分布を算出するステップを含むことができ、粗視化経験確率分布、粗視化混合確率分布、および粗視化クラス帰属確率に基づき、評価関数を算出することができる。

【0043】

このような粗視化を行うことにより、処理時間を大幅に短縮することができる。以下、粗視化処理を詳細に説明する。

【0044】

本発明の染色体テリトリの抽出方法としては、以下のものも有効である。

(1) 抽出の対象となる像を含む画像に対し、画像上の各ピクセルを複数のクラスに分類し、所望の領域を抽出する抽出方法において、

画像の各ピクセルが取り得る属性値全体から構成されるデータ空間を、所与の解像度で部分空間に分解し、該各部分空間に属性値を取るピクセルの集合と、該ピクセルの属性値の平均と、該ピクセルの個数とを保持する粗視化データ空間を構成する第1の工程と、

各部分空間のピクセル数を画像に含まれる全ピクセル数で除し、粗視化データ空間上の粗視化経験確率分布を算出する第2の工程と、

各クラスの属性を規定するクラスパラメータおよびクラスの混合比率を初期化する第3の工程と、

クラスパラメータからクラスが指定されたもとの条件付き確率分布を算出し、当該条件付き確率分布を各部分空間内で平均化することによって粗視化条件付き確率分布を算出する第4の工程と、

クラスの混合比率に粗視化条件付き確率分布を乗じることにより、各ピクセルが各クラスに帰属する確率であるクラス帰属確率を算出する第5の工程と、
クラスパラメータおよびクラスの混合比率を更新する第6の工程と、
粗視化条件付き確率分布を用いて、評価関数を算出する第7の工程と、
評価関数が所与の終了条件を満たしているか否かを調べる第8の工程と、
評価関数が所与の終了条件を満たしたときのクラスパラメータ、クラスの混合比率、およびクラス帰属確率に基づいて、各ピクセルが属するクラスを決定し、所望の領域を抽出する第9の工程と、
を含むことを特徴とする抽出方法。

【0045】

ここで、評価関数としては、粗視化対数尤度を用いることができる。このとき、第6の工程において、評価関数を増加させるようにクラスパラメータおよびクラスの混合比率を更新することができる。また、第3の工程において、クラスの数を実初期化することもできる。また、クラス数は、予め所与の値を設定しておくこともできる。また、本発明の抽出方法において、第8の工程において、評価関数が所与の条件を満たすまで、第4、第5、第6、第7、および第8の工程を繰り返すことができる。

【0046】

(2) 上記(1)に記載の抽出方法において、

第4の工程において、粗視化条件付き確率分布を算出する際に、各部分空間に含まれるピクセルの属性値の平均値を算出し、該平均値を用いて各部分空間における粗視化条件付き確率分布を算出することを特徴とする抽出方法。

【0047】

(3) 上記(1)または(2)に記載の抽出方法において、

粗視化の解像度が元の解像度であるかを調べる第10の工程と、

部分空間の解像度を元の解像度へ復元する第11の工程と、をさらに含み、

部分空間の解像度を元の解像度へ復元した後に、第1の工程～第8の工程を行い、

第3の工程において、第9の工程におけるクラスパラメータおよびクラスの混合比率を実初期値として用いることを特徴とする抽出方法。

【0048】

(4) 上記(1)乃至(3)いずれかに記載の抽出方法において、

第9の工程において、前記評価関数が所与の終了条件を満たしたときのクラスの混合比率に画像に含まれる全ピクセル数を乗じることにより各クラスに属するピクセル数を算出し、クラス帰属確率が高い順から前記ピクセル数のピクセルを選択することにより各クラスに属するピクセルを決定することを特徴とする抽出方法。

【0049】

(5) 上記(1)乃至(4)いずれかに記載の抽出方法において、

第7の工程において、評価関数としてAICを用い、第6の工程において評価関数を減少させるようパラメータを変更することを特徴とする抽出方法。

【0050】

(6) 上記(1)乃至(4)いずれかに記載の抽出方法において、

第7の工程において、評価関数としてMDL (Minimum Description Length, 最小記述長)を用い、第6の工程において該評価関数を減少させるようパラメータを変更することを特徴とする抽出方法。

【0051】

(7) 上記(1)乃至(4)いずれかに記載の抽出方法において、

第7の工程において、評価関数としてストラクチャル・リスク (Structural Risk)を用い、第6の工程において該評価関数を減少させるようパラメータを変更することを特徴とする抽出方法。

【0052】

(8) 上記(1)乃至(7)いずれかに記載の抽出方法において、

第3の工程において、

各部分空間どうしが近傍にあるか否かを定義する近傍半径と、クラスの数を設定し（ステップB1）、

各部分空間ごとに、各部分空間の代表値を設定し（ステップB2）、

分類対象部分空間の集合を設定し（ステップB3）、

分類対象部分空間の集合の中から、粗視化経験確率の最も高い部分空間を選択し（ステップB4）、

粗視化経験確率の最も高い部分空間の代表値との距離が、近傍半径以内である代表値を持つ部分空間すべてを、近傍集合として選択し（ステップB5）、

すでに分類が完了したクラスに含まれる部分空間の代表値と、近傍集合に含まれる部分空間の代表値との間の最短距離が、近傍半径よりも大きいかな否かを調べ（ステップB6）、

ステップB6において、すでに分類が完了したクラスに含まれる部分空間の代表値と近傍集合に含まれる部分空間の代表値との間の最短距離が近傍半径よりも大きければ、近傍集合を新規クラスとするとともに分類対象部分空間から近傍集合を削除し（ステップB7）、

ステップB6において、すでに分類が完了したクラスに含まれる部分空間の代表値と近傍集合に含まれる部分空間の代表値との間の最短距離が近傍半径より小さければ、近傍集合をすでに分類済みのクラスに追加するとともに分類対象部分空間から近傍集合を削除し（ステップB8）、

分類対象部分空間が空集合であるかな否かを調べ（ステップB9）、

ステップB9において、分類対象部分空間が空集合でなければ、ステップB4以下を繰り返す、

ステップB9において、分類対象部分空間が空集合であれば、分類が完了したクラスの数に所与の数以上あるかな否かを調べ（ステップB10）、

ステップB10において、分類が完了したクラスの数に所与の数よりも少なければ、近傍半径を縮小し（ステップB11）、ステップB3以下を繰り返す、

ステップB10において、分類対象部分空間が空集合であり、かつ分類済みのクラスの数に所与の数よりも多ければ、各クラス内でクラスパラメータを算出し、これをクラスパラメータの初期値とするとともに、各クラス内に含まれる部分空間の数の比率をクラスの混合比率の初期値とする（ステップB12）、
ことを特徴とする抽出方法。

【0053】

なお、ステップB7およびB8において、すでに分類が完了したクラスに含まれる部分空間の代表値と近傍集合に含まれる部分空間の代表値との間の最短距離が近傍半径と等しい場合、いずれの処理を行ってもよい。

【0054】

(9) 抽出の対象となる像を含む画像に対し、画像上の各ピクセルを複数のクラスに分類し、所望の領域を抽出する抽出装置において、

画像を読み込む入力装置と、

画像の各ピクセルが取り得る属性値全体から構成されるデータ空間を、所与の解像度で部分空間に分解し、該各部分空間に属性値を取るピクセルの集合と、該ピクセルの属性値の平均と、該ピクセルの個数を保持し、粗視化データ空間を構成する領域粗視化装置と、

各部分空間のピクセル数を画像に含まれる全ピクセル数で割り、粗視化データ空間上の粗視化経験分布を算出する粗視化経験確率分布算出装置と、

各クラスの属性を規定するクラスパラメータおよびクラスの混合比率を初期化し、各クラスの属性を規定するクラスパラメータから、クラスが指定されたもとでの条件付き確率分布を算出し、クラスが指定されたもとでの条件付き確率分布を、各部分空間内で平均化することによって粗視化条件付き確率分布を算出する粗視化条件付き確率分布装置と、

粗視化条件付き確率分布から画像各ピクセルが各クラスに帰属する確率であるクラス帰

属確率を算出するクラス帰属確率算出装置と、

クラスパラメータおよびクラスの混合比率を更新するパラメータ更新装置と、

粗視化条件付き確率分布を用いて、評価関数を算出する評価関数算出装置と、

評価関数が所与の終了条件を満たしているか否かを調べ、評価関数が所与の終了条件を満たしたときのクラスパラメータ、クラスの混合比率、およびクラス帰属確率に基づいて、各ピクセルが属するクラスを決定し、所望の領域を抽出する領域抽出装置と、

抽出した領域を出力する出力装置と、

を有することを特徴とする抽出装置。

【0055】

ここで、評価関数算出装置は、評価関数として、粗視化対数尤度、AIC、MDL、またはストラクチャル・リスクを用いることができる。評価関数算出装置が評価関数として粗視化対数尤度を用いるとき、パラメータ更新装置は、評価関数を増加させるように、クラスパラメータおよびクラスの混合比率を更新することができる。また、評価関数算出装置が評価関数としてAIC、MDL、またはストラクチャル・リスクのいずれかを用いるとき、パラメータ更新装置は、評価関数を減少させるようにクラスパラメータおよびクラスの混合比率を更新することができる。

【0056】

(10) 上記(9)に記載の抽出装置において、

評価関数が所与の終了条件を満たしていることが確認されたのち、粗視化の解像度が元の解像度であるかを調べ、データ空間の解像度を元の解像度へ復元する解像度復元装置をさらに有することを特徴とする抽出装置。

【0057】

以下、上記(1)乃至(10)に記載の抽出方法および抽出装置の作用を説明する。

上記(1)に記載の抽出方法においては、画像を構成する各ピクセルの持つ属性値を確率変数と見なし、推定したピクセル値の確率分布に基づいて所望の領域を抽出する。ここで、属性値としては、たとえばモノクロ画像ならば輝度の値、カラー画像ならば赤(R)、緑(G)、青(B)の色要素の強度等を用いることができる。

【0058】

ここで、所望の領域を抽出するために、各ピクセルを、各ピクセルの属性値に基づいて類似の属性を持つ複数のグループに分類する。本明細書においては、類似の属性を持つピクセルの集合をクラスと呼ぶ。各クラスは、たとえばそのクラスに属する属性値の平均値や分散等によって特徴づけられるものとする。以下、これらクラスの特徴をそのクラスの「クラスパラメータ」とよび、 ϕ_i ($i=1, \dots, k$) で表す。ここで、 k はクラスの数である。ここで、 j 番目のピクセルが、 x_j という属性値を取る確率は、次の混合分布で表現できる。

【0059】

【数1】

$$p(x_j) = \sum_{i=1}^k w_i f(x_j | \phi_i) \quad \dots\dots(1)$$

【0060】

ここで、 $f(x_j | \phi_i)$ は、データが i 番目のクラスから発生していると仮定したときの条件付確率分布、 w_i は各クラスの混合比率で、

【0061】

【数 2】

$$\sum_{i=1}^k \omega_i = 1$$

を満たす。

【0062】

画像がモノクロ画像の場合、 x_j は、0 から 255 等の整数値で表される。また、画像がカラー画像の場合、 x_j は、RGB の色要素の値を成分とした 3 次元ベクトル (x_{j1} , x_{j2} , x_{j3}) で表される。ここで、各 x_{j1} ($1=1, 2, 3$) は、たとえば 0 から 255 の整数値を取る。

【0063】

混合比率 w_i は、異なるクラスに属する領域の面積比 (2 次元領域の場合) または体積比 (3 次元領域の場合) を表す。たとえば、平均輝度 200、輝度の標準偏差が 20 という特徴で表される明るい画像領域 (クラス 1 とする) と、平均輝度 50、輝度の標準偏差が 10 という特徴で表される暗い領域 (クラス 2 とする) の二つの領域から構成されるモノクロ画像があるとする。また、その画像の面積の 7 割を明るい領域が、3 割を暗い領域が占めているとする。この場合、クラスの数 $k=2$ 、クラスパラメータは $\phi_1 = (200, 20)$, $\phi_2 = (50, 10)$ であり、この画像の混合分布は、

【0064】

【数 3】

$$p(x_j) = 0.7(x_j|200, 20) + 0.3(x_j|50, 10) \quad \dots\dots (2)$$

と表現することができる。

【0065】

以下、クラスの混合比率 w_i とクラスパラメータ ϕ_i をまとめて θ_i と示す。以下において単に「パラメータ」というときはこの θ_i を意味する。

【0066】

上記 (1) に記載の抽出方法では、以下のように定義される平均対数尤度

【0067】

【数 4】

$$L = \frac{1}{n} \sum_{j=1}^n \log p(x_j) = \frac{1}{n} \sum_{j=1}^n \log \left[\sum_{i=1}^k w_i f(x_j|\phi_i) \right] \quad \dots\dots (3)$$

を最大にするパラメータを推定し、推定したパラメータの情報を利用して領域抽出を行う。ここで n は画像に含まれるピクセル数である。このような統計手法は最尤法と呼ばれている。しかしながら、平均対数尤度を最大にするパラメータを推定することは、一般に困難である。そこで、この平均対数尤度のかわりに次の量

【0068】

【数 5】

$$Q = \frac{1}{n} \sum_{j=1}^n \sum_{i=1}^k \pi_{ij} \log [w_i f(x_j|\phi_i)] \quad \dots\dots (4)$$

で表される完全対数尤度の期待値 Q を用いてパラメータを推定することができる。

ここで、
 【0069】
 【数6】

$$\pi_{ij} = \frac{w_i f(x_j | \phi_i)}{\sum_{\ell=1}^k w_{\ell} f(x_j | \phi_{\ell})} \quad \dots\dots (5)$$

は、j 番目のピクセルが i 番目のクラスに帰属する確率である。本発明ではこれをクラス帰属確率という。Qが増大するようにパラメータを更新すれば、上述の平均対数尤度Lも必ず増大することは、数学的に証明されている（たとえばA. P. Dempster, N. M. Laird, and D. B. Rubin, Maximum Likelihood From Incomplete Data via The EM Algorithm, J. Roy. Stat. Soc., vol. 30, pp. 205-248, 1977（以下、非特許文献3という））。

【0070】

本発明において、パラメータを推定し、推定されたパラメータから実際に領域を抽出する手順を説明する。

まず適当な初期パラメータから開始し、式（5）で与えられるクラス帰属確率を求める。次にQを増加させるようにパラメータwおよび ϕ を更新し、改めてQを算出する。最終的にQが増加しなくなるまでこの手順を繰り返す。最終的にQが増加しなくなったときのパラメータwおよび ϕ を用いて領域を抽出する。k個のクラスのうち、たとえばi番目のクラスに属するピクセルを抽出するためには、まず各ピクセルのi番目のクラスへの帰属確率の値を調べる。ついで、その確率値が一定値以上のピクセルをそのクラスに属するものとして分類する。すべてのピクセルをそれぞれのクラスに分類し終わった時点で、k個のクラスの中から所望の属性を持ったクラスを選択し、そのクラスに属するピクセルを画像から抽出すれば、所望の領域を自動的に抽出することができる。

【0071】

本発明では、このQの最大化を高速に行うために、粗視化確率分布を導入する。粗視化確率分布は、データが取り得る値全体からなる空間（以下、データ空間と呼ぶ）を互いに交わらないN個の部分空間に分解し、各部分空間に確率を割り当てることによって構成される。

【0072】

具体的には、j 番目の部分空間における粗視化された条件付き確率分布を

【0073】

【数7】

$$\bar{f}_j(\phi_i) = \frac{1}{m(A_j)} \int_{A_j} f(x | \phi_i) dx \quad \dots\dots (6)$$

で定義し、粗視化条件付き確率分布を

【0074】

【数8】

$$\bar{f}(x | \phi_i) = \sum_{j=1}^N \bar{f}_j(\phi_i) I_{A_j}(x) \quad \dots\dots (7)$$

とする。ここで、 A_j は j 番目の部分空間である。Dをデータ空間全体とすると、

【0075】

【数9】

$$\mathcal{D} = \cup_j A_j, \quad A_i \cap A_j = \emptyset \quad (i \neq j) \quad \dots\dots (8)$$

を満たす。また $I_A(x)$ は、データ値が部分空間 A に含まれるときに 1、それ以外の場合に 0 となる指示関数で、

【0076】

【数10】

$$m(A) = \int_A dx$$

は、 A の測度である（データ空間が 2 次元のときには A の面積を、3 次元空間のときは体積を表す。）。

【0077】

このように定義された粗視化条件付き確率分布を用いると、上記式 (4) に示した Q は、

【0078】

【数11】

$$\bar{Q} = \sum_{i=1}^n \bar{q}_i \pi_{ij} \log [w_i \tilde{f}_j(\phi_i)] \quad \dots\dots (9)$$

と書き換えることができる。ここで

【0079】

【数12】

$$\bar{q}_j = \frac{1}{n} \sum_{\ell=1}^n I_{A_j}(x_\ell) \quad \dots\dots (10)$$

は粗視化された経験確率分布、

【0080】

【数13】

$$\pi_{ij} = \frac{w_i \tilde{f}_j(\phi_i)}{\sum_{i=1}^k w_i \tilde{f}_j(\phi_i)} \quad \dots\dots (11)$$

は粗視化されたクラス帰属確率である。この式 (9) で与えられる粗視化された完全対数尤度を最大化することによって、次の粗視化平均対数尤度

【0081】

【数 14】

$$\tilde{L} = \sum_{j=1}^N \tilde{g}_j \log \left[\sum_{i=1}^k w_i \tilde{f}_j(\phi_i) \right] \quad \dots\dots (12)$$

を最大化することができる。

【0082】

元のQと比較すると、式(4)で与えられるQが全データに関して和をとっているのに対し、式(9)で与えられる粗視化された完全対数尤度は、部分空間に関してのみ和を取っている。後に示すように、このような粗視化を行うことにより計算量を大幅に軽減することができる。たとえば、512×512ピクセルの画像の場合、式(4)を用いると、26万個以上のデータについて和を取る必要があるが、本発明の粗視化分布を用いた場合、部分空間に関する和を1000個程度にまで軽減することができ、高速な推定を行うことができる。

【0083】

さらに、上記(2)に記載の抽出方法では、各部分空間における粗視化された確率値を、その部分空間に含まれるデータの平均値における確率値で近似する。

【0084】

【数 15】

$$\tilde{f}_j(\phi_i) = f(\tilde{x}_j | \phi_i) \quad \dots\dots (13)$$

ここで、

【0085】

【数 16】

$$\tilde{x}_j = \frac{1}{\sum_{\ell=1}^n I_{A_j}(x_\ell)} \sum_{\ell=1}^n x_\ell I_{A_j}(x_\ell) \quad \dots\dots (14)$$

は、j部分空間 A_j に含まれるデータの平均値である。この近似によって、部分空間内における積分（あるいは総和）操作を省略することができ、さらに計算量を軽減することができる。

【0086】

さらに、粗視化確率分布を用いて推定したパラメータを初期値とし、元の解像度で再度推定を行うこともできる。この場合、粗視化確率分布を用いてすでにほぼ最適なパラメータが得られているため、最初から元の解像度で推定を行う場合に比較して、パラメータの逐次更新の回数がかかるかに少なくすむため、高精度な推定を高速に行うことができる。

【0087】

さらに、領域を抽出する際、推定された混合比率 w_i に全ピクセル数を掛け、i番目の領域に属するピクセル数を推定する処理を行うこともできる。そして領域帰属確率の高い順に上位 n_i 個をこの領域に属するピクセルとして抽出する。この方法により、どの値までの確率値をその領域に属しているものと見なすか、という閾値を自動的に決定することができる。本発明において、染色体テリトリの形状状態を定量化し、その定量化結果に基づいて細胞の状態を評価するためには、画像から染色体テリトリを精度良く抽出する必要がある。このように閾値を自動的に決定することにより、染色体テリトリを精度良く抽出することができる。

【0088】

さらに、評価関数としては、それぞれAIC、MDL、ストラクチャル・リスクを用い、その値がもっとも小さい結果を与えるモデルを選択することもできる。評価関数としてAIC、MDL、またはストラクチャル・リスクを用いた場合、過剰なパラメータ数を用いると評価関数の値が増加してしまうため、最適なパラメータ数を推定することが可能となる。これにより、画像が何種類の領域から形成されているかの適切な領域数を推定することができる。

【0089】

以上のように、画像から染色体テリトリを抽出する際に、粗視化処理を行うと、粗視化処理を行わない場合に比べて、条件付き確率分布やクラス帰属確率の計算量を大幅に削減することができるため、染色体テリトリの抽出に要する時間を大幅に短縮することができる。このような処理の結果推定されるクラスパラメータは、粗視化処理を行わない場合に比べて精度が劣ることになるが、本発明の抽出方法では、クラスパラメータから算出されるクラス帰属確率に基づいて領域抽出を行うため、粗視化による誤差の影響をうけることなく精度良く領域抽出を行うことが可能である。

【0090】

以上、本発明の構成について説明したが、これらの構成を任意に組み合わせたものも本発明の態様として有効である。また、本発明の表現を他のカテゴリーに変換したものもまた本発明の態様として有効である。

【発明の効果】

【0091】

本発明の評価方法によれば、細胞周期の間期における染色体テリトリの配置状態を定量化することができる。これにより、細胞の状態を統計的に評価することもできる。

【発明を実施するための最良の形態】

【0092】

図1は、本発明の実施の形態における、細胞の状態の評価方法の処理手順を示すフローチャートである。

まず、染色体テリトリを含む細胞を撮影した画像から、染色体テリトリの像を抽出する(S20)。本実施の形態において、3次元画像から細胞の核および複数の染色体テリトリの像をそれぞれ抽出する。その後、染色体テリトリの配置状態を定量化する(S22)。染色体テリトリの配置状態を定量化した結果は保存される。複数の検体について、ステップ20およびステップ22の処理を行い、複数の検体の定量化結果を蓄積する(S24)。また、ある検体について、ステップ20およびステップ22の処理を行い、定量化した結果をステップ24で蓄積された複数の検体の定量化結果と比較して、その検体における細胞状態を評価する(S26)。

【0093】

図2は、図1に示した処理手順を実行する本実施の形態における評価装置を示すブロック図である。

評価装置100は、画像取得部102と、抽出処理部104と、定量化処理部106と、評価部110と、出力部112、と、定量化結果記憶部114とを含む。画像取得部102は、染色体テリトリを含む細胞を撮影した3次元画像を取得する。抽出処理部104は、3次元画像から染色体テリトリの像を抽出する。定量化処理部106は、染色体テリトリの配置状態を定量化する。定量化処理部106は、定量化した結果を定量化結果記憶部114に保存する。評価部110は、ある検体の定量化結果を定量化結果記憶部114に記憶された複数の検体の定量化結果と比較して、その検体の細胞状態を評価する。出力部112は、抽出処理部104による染色体領域の像の抽出結果や、評価部110による細胞状態の評価結果を出力する。出力部112は、たとえばディスプレイ装置やプリンタ等により構成することができる。

【0094】

以下、図1に示した各手順を説明する。

(染色体テリトリの抽出 (S20))

染色体テリトリを含む細胞の3次元画像は、3D-FISH (Fluorescent in situ hybridization) 法を用いて得ることができる。画像取得部102は、画像スキャナやデジタルカメラ等により構成することができ、3D-FISH法を用いて得られた画像を読み取ることができる。また、画像取得部102は、生物顕微鏡に接続されたコンピュータにより構成することもでき、ネットワークを通して画像を直接入力することも可能である。

【0095】

以下、3D-FISH法の手順を説明する。

カバーガラス上で細胞を培養し、PFA (Paraformaldehyde) で細胞を固定する。つづいて、Triton X100溶液で処理を行う。次いで、液体窒素を使用し、凍結融解を数回繰り返す。その後、塩酸処理およびペプシン処理を行う。これにより、細胞膜および核膜が部分的に破壊され、プローブの浸透性を高めることができる。また、以上の処理により、核以外の部分にある程度除去することができる。

【0096】

つづいて、細胞を再度固定し、約75℃でDNAを変性させる。これにより、DNAの二本鎖が一本鎖となる。ここへ、標識したプローブ (目的の染色体テリトリ、および核と相補的なDNA断片) を加え、約37℃で反応させると、プローブが目的の染色体テリトリおよび核に相補的に結合する。プローブには、目的の染色体テリトリの種類および核毎に異なる波長によって発色する色素が付けられている。これにより、特定の色素が発色する波長を照射することにより、その色素のみを特異的に発色させることができる。

【0097】

以上のようにして準備した試料に共焦心レーザ顕微鏡を用いて特定の波長を照射し、目的の染色体テリトリを含む画像を得る。各染色体テリトリを含む画像、および核を含む画像を重ね合わせることににより、複数の染色体テリトリを含む画像を得ることができる。

以下、このようにして得られた画像に施す処理を説明する。以下、単に「染色体テリトリ」、「核」というときは、染色体テリトリや核そのものではなく、画像上における染色体テリトリ像および核の像を示す。

【0098】

図2に示した抽出処理部104は、以上のようにして得られた画像から、染色体テリトリを抽出する画像処理を行う。本実施の形態において、抽出処理部104は、複数のピクセルを含む画像を複数のクラスに分類し、複数のクラスをそれぞれ染色体テリトリおよび核に対応付けることににより、画像から染色体テリトリおよび核の像を抽出する。この際、抽出処理部104は、各ピクセルに輝度が異なる複数の輝度源の存在を仮定し、各輝度源がそのピクセルに属する確率、および画像における輝度源の混合比率に基づいてクラス分類を行う。

【0099】

このような手法は、たとえば特開2001-92980号公報 (以下特許文献1という。) に記載されている。特許文献1においては、画像データから、画像上の各ピクセルが領域の各々に属する領域帰属確率の期待値を算出し、領域帰属確率と領域パラメータから求められる混合確率分布に基づいて評価関数を算出し、そのときの領域帰属確率に基づいて各領域を分離し、分離された領域に基づいて輪郭を抽出する手法が記載されている。本実施の形態においては、この手法にさらに粗視化の概念を導入し、最適なパラメータを高速に求める。

【0100】

図3は、図2に示した抽出処理部104を示すブロック図である。図4は、抽出処理部104の処理手順を示すフローチャートである。

【0101】

本発明の実施の形態における抽出処理部104は、画像データを読み込む入力装置1とデータ空間粗視化装置2と、パラメータ更新装置6と、粗視化経験確率分布算出装置3と、粗視化条件付き確率分布算出装置4と、クラス帰属確率算出装置5と、評価関数算出装

置 7 と、領域抽出装置 8 と、出力装置 9 とを含む。

【0102】

入力装置 1 は、上述したような 3D-FISH 法により得られた細胞のカラー画像を画像取得部 102 から入力する（ステップ A1）。入力装置 1 は、読み込んだデータをデータ空間粗視化装置 2 に送る。

【0103】

データ空間粗視化装置 2 は、データ空間を互いに交わらない N 個の部分空間に分解し、粗視化データ空間を構成する（ステップ A2）。ここで、データ空間とは、画像の各ピクセルが取り得る属性値全体の集合を意味する。たとえば標準的グレースケールの画像において、各ピクセルの属性は輝度によって表すことができ、輝度は通常 0 から 255 の整数値で表現される 1 次元空間である。また、標準的なカラー画像の場合ならば、R、G、B 各色要素に対し、通常 0 から 255 の整数値で表現される 3 次元空間となる。

【0104】

データ空間粗視化装置 2 は、たとえば解像度を 8 とすると、RGB の各値を 8 ずつに区切り、 $8 \times 8 \times 8$ の立方体を 1 つの部分空間とすることができる。解像度は 3 次元の軸で同一でなくともよい。たとえば R 軸を h_1 ずつ、G 軸を h_2 ずつ、B 軸を h_3 ずつに区切って粗視化してもよく、各部分空間が重なりを持たず、かつ全データ空間をもれなく覆うことができればよい。以下において、解像度 h で粗視化する、と表現する場合は、 $h \times h \times h$ で RGB の各値を区切ることを意味する。

【0105】

また、データ空間粗視化装置 2 は、各部分空間に値を取るピクセルの集合（ピクセルの番号）と、それらのピクセルの属性値の平均と、ピクセルの個数を保持する。

【0106】

図 5 は、解像度 $h_1 \times h_2 \times h_3$ で粗視化したときの粗視化データ空間 11、および部分空間 12 を示す。

たとえば j 番目のピクセルの属性値 x_j が、 $R=128$ 、 $G=255$ 、 $B=18$ である場合、解像度 8 で粗視化すると、このピクセルは、(16、31、2) なる指標で指定される部分空間に値を取る。このことを以下では、 x_j がこの部分空間に含まれる、と表現する。

【0107】

図 3 に戻り、データ空間粗視化装置 2 は、粗視化データ空間を構成した後、各部分空間に含まれるピクセルの番号と、ピクセルの個数を粗視化経験確率分布算出装置 3 に送る。

【0108】

粗視化経験確率分布算出装置 3 は、粗視化経験確率分布を算出する（ステップ A3）。ここで粗視化経験確率分布とは、粗視化データ空間の各部分空間に含まれる観測データを、全ピクセル数で割った値から算出された確率分布を表す。粗視化経験確率分布算出装置 3 は、粗視化経験確率分布を粗視化条件付き確率分布算出装置 4 に送る。

【0109】

粗視化条件付き確率分布算出装置 4 は、クラスパラメータを初期化する（ステップ A4）。パラメータの初期値の決定手順は後述する。また、粗視化条件付き確率分布算出装置 4 は、粗視化条件付き確率分布を算出する（ステップ A5）。粗視化条件付き確率分布は、具体的には、以下のように算出する。ここでは、 j 番目のピクセル値が i 番目のクラスから発生しているとした条件のもとでの条件付き確率が、以下のような多次元正規分布

【0110】

【数 17】

$$f(x|\phi_i) = \frac{1}{\sqrt{(2\pi)^3 |\Sigma_i|}} \exp \left[-\frac{1}{2} (x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i) \right] \dots\dots (15)$$

で与えられているとする。ここで、 x は、RGB 各色の値からなる 3 次元ベクトル、 μ_i

は i 番目のクラスの平均色を表す 3 次元ベクトル、 Σ_i は i 番目のクラスの共分散行列であり、 $|\Sigma_i|$ 、 Σ_i^{-1} はそれぞれ行列 Σ_i の行列式、逆行列を表す。また、 $(x - \mu_i)^T$ は転置を表す。

画像がグレースケールの場合、 j 番目のピクセル値が i 番目のクラスから発生しているとした条件のもとでの条件付き確率は、以下のような一次元正規分布

【数 18】

$$f(x|\phi_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp[-(x - \mu_i)^2 / 2\sigma_i^2]$$

で与えられる。ここにおける x は輝度の値を示すスカラー値、 μ_i は i 番目のクラスの平均輝度を表すスカラー値、 σ_i は i 番目のクラスの輝度の標準偏差である。

【0111】

粗視化条件付き確率分布算出装置 4 は、この条件付き確率を式 (6) で与えられる式で計算する。このとき、各部分空間の測度 $m(A_j)$ は、各部分空間の体積となる。たとえば、解像度 8 で一様に粗視化した場合、 $8 \times 8 \times 8 = 512$ となる。

【0112】

粗視化条件付き確率分布算出装置 4 は、粗視化条件付き確率分布を式 (13) によって近似することもできる。この方法により、パラメータを更新するたびに式 (6) で与えられる演算をする必要がなくなり、計算量を大幅に削減することができる。

【0113】

粗視化条件付き確率分布算出装置 4 は、粗視化経験確率分布に基づいてパラメータの初期値を決定することも可能である。各部分空間を大雑把にクラス分類し、分類の結果得られた各クラス内で平均値や分散値を求め、これらの値をパラメータ推定の初期値とする。

【0114】

図 6 は、粗視化経験分布に基づいて、パラメータの初期値を決定する処理手順を示したフローチャートである。以下、図 6 を参照して、パラメータの初期値を設定する手順を説明する。

【0115】

まず、ステップ B1 において、近傍半径と、分類すべきクラスの個数とを設定する。近傍半径とは、その範囲内にあるすべての部分空間を同じクラスに属するとみなし、大雑把にクラス分類をするための基準値のことである。たとえばカラー画像の場合、同じような色のピクセルどうしは RGB 値も近く、したがって同じクラスに分類するのが自然であると考えられるからである。

【0116】

近傍半径が大きすぎる場合、所望のクラス数に到達する前に分類が完了してしまう場合がある。その場合、後述するように、近傍半径を縮小して再度クラス分類を行う。そのため、最終的には必要な数のクラスに分類することができる。したがってこの近傍半径の初期値は十分大きな値、たとえば 50 等に設定する。分類すべきクラスの個数は、所与の値をそのまま用いる。

【0117】

次に、ステップ B2 において、各部分空間ごとに、各部分空間の代表値を設定する。各部分空間の代表値としては、たとえば部分空間の中央値等を用いることができる。以下では、これら代表値間の距離を部分空間の距離とする。

【0118】

次に分類対象となる部分空間の集合を設定する (ステップ B3)。以下ではこの集合を Ω と書く。 Ω の初期値は、データを含む全部分空間全体から成る集合とする。また分類済みクラスの数 i を 1 とし、分類済みクラス C_i の初期値を空集合とする。

【0119】

次に、 Ω に属する部分空間の中から、粗視化経験確率の最も高い部分空間を選択する (ステップ B4)。この部分空間を A_s と書く。次に、 Ω に属する部分空間と A_s との距離

を調べ、近傍半径 r 以内にある部分空間をすべて選択し、これを近傍集合とする（ステップ B5）。以下では、近傍集合を B_s と書く。

【0120】

次に、すでに分類が完了しているクラス C_i に含まれる部分空間と、近傍集合 B_s に含まれる部分空間との最短距離を求め、近傍半径 r よりも大きいか否かを調べる（ステップ B6）。もしこの最短距離が r よりも大きいならば、近傍集合 B_s は、すでに分類が完了したクラスと十分異なる属性を持ち、かつ高い確率で出現しているから、新しいクラスと見なしてよい。したがって、近傍集合をそのまま新規クラスとして採用する。 B_s は分類が完了したので、分類対象集合 Ω から削除する。図6においてはこの削除を、記号を用いて「 $\Omega \leftarrow \Omega \setminus B_s$ 」と示す。 Ω を更新した後、ステップ B4 に戻る（ステップ B7）。

【0121】

すでに分類が完了しているクラス C_i に含まれる部分空間と、近傍集合 B_s に含まれる部分空間との最短距離が、近傍半径 r よりも小さければ、近傍集合 B_s は C_i と属性が近いと考えられるため、 B_s を C_i に統合する。 B_s は分類が完了したので、分類対象集合 Ω から削除する（ステップ B8）。

【0122】

次に、 Ω が空集合であるか否かを調べ（ステップ B9）、空集合でなければ、ステップ B4 へ進み、 Ω が空集合であれば、分類が完了したクラスの数 k が k 個以上あるか否かを調べ（ステップ B10）、 k 個以下なら近傍半径に 1 より小さい定数をかけて近傍半径を縮小する。この定数としては、たとえば 0.9 等の値を用いることができる。その後、ステップ B3 以下を繰り返す。

【0123】

ステップ B9 において、 Ω が空集合であり、かつ分類済みのクラスの数 k が所与の数よりも多ければ、所望の数のクラス分類が完了しているので、各クラス内でクラスパラメータを算出し、これをクラスパラメータの初期値とし、また各クラス内に含まれる部分空間の数の比率をクラスの混合比率の初期値とする（ステップ B12）。

【0124】

図7は、図6を参照して説明した手順を具体的に説明する図である。ここではデータ空間が一次元であるとし、全部で 10 個の部分空間が設定されていると仮定する（図7（a））。図7（a）において、横軸は部分空間の番号を表し、縦軸は粗視化経験確率分布を表す。以下では処理の流れを直観的に説明することに主眼を置くため、粗視化経験確率値や部分空間の代表値、近傍半径については具体的な数値は用いない。

【0125】

ステップ B1 において、たとえばクラス数を 2 とし、近傍半径を r とする。ステップ B2 において、各部分空間の代表値を設定する。ステップ B3 において、分類対象集合の Ω 初期値は、データを含んだ全部分空間とするので、

$$\Omega = \{A_3, A_4, A_5, A_6, A_7, A_8\}$$

となる。 A_1 、 A_2 、 A_9 、および A_{10} は、粗視化確率が 0、すなわちこれらの部分空間に含まれるデータは観測されていないので、分類対象集合には含めない。

【0126】

ステップ B4 で、分類対象集合に含まれる部分空間の中で、最も高い粗視化経験確率を持つ A_7 を選び、これを A_s とする（図7（a））。ステップ B5 で、 A_7 から近傍半径 r 内にある部分空間を選択し、これを B_s とする。図7（a）に示された近傍半径内にある部分空間は、 A_5 、 A_6 、 A_7 、および A_8 であるから、

$$B_s = \{A_5, A_6, A_7, A_8\}$$

となる（図7（b））。

【0127】

ステップ B6 において、分類されているクラスはまだないので、この B_s をそのまま最初のクラス C_1 として採用し、分類対象集合から B_1 を取り除き、ステップ B4 に戻る。図7（b）において、粗視化経験確率の高さを示す棒グラフが白抜きで表されているのは

、部分空間が分類対象集合から除かれたことを表す。

【0128】

ステップB4において、残りの分類対象集合の中で最も高い粗視化経験確率を持つものは、 A_4 であるから、これを新たに A_5 とする(図7(c))。ステップB5で、 A_4 から近傍半径 r 内にある部分空間を選択して、これを B_5 とする。ここでは、

$B_5 = \{A_3, A_4\}$ となる。

【0129】

ステップB6で、分類済みのクラス、すなわち

$C_1 = \{A_5, A_6, A_7, A_8\}$

を調べると、 A_4 から近傍半径 r 内にある A_5 および A_6 を含んでいる。したがって現在の B_5 を分類済みのクラス C_1 に統合する(図7(d))。

【0130】

これで分類対象集合は空となり、すべての部分空間の分類が完了したことになるが、分類されたクラス数は1であり、所望のクラス数2に達していない(ステップB10)。したがって近傍半径を縮小し(ステップB11)、ステップB3以下を繰り返す。

【0131】

以下、縮小された半径を r' とし(図7(e))、上述の説明と同様の手続きを繰り返す。ただし、今回は近傍半径が縮小されているため、以下のような違いが生じる。すなわち、 A_7 の近傍半径 r' 内にある部分空間は、今回の場合

$B_5 = \{A_6, A_7, A_8\}$

となる(図7(f))。

この B_5 をそのまま最初のクラス C_1 として採用し、残りの分類対象集合の中から最も高い粗視化経験確率を持つ A_4 を選ぶ(図7(g))。 A_4 から近傍半径 r' 内にある部分空間は、

$B_5 = \{A_3, A_4, A_5\}$

となる。今回は、分類済みのクラス C_1 の中に、 A_4 から近傍半径 r' 内にある部分空間は存在しないので、現在の B_5 を新規クラス C_2 として採用する(図7(h))。これですべての部分空間が所望の二つのクラスに分類されたことになる。

【0132】

大雑把なクラス分類が完了すれば、分類されたクラス内で平均や分散を求め、以降で行う推定の初期パラメータとすることができる。このような方法でパラメータの初期値を設定することにより、この後の計算量を削減して領域抽出を高速に行うことができる。また、

初期パラメータを適切に設定することは、粗視化対数尤度の最大化の過程で局所最適解に陥ることを防止する上で有効である。

【0133】

図3に戻り、粗視化条件付き確率分布算出装置4は、このようにして求めたパラメータを初期値として粗視化条件付き確率分布を求めることができる。粗視化条件付き確率分布算出装置4は、求めた粗視化条件付き確率分布をクラス帰属確率算出装置5に送る。

【0134】

クラス帰属確率算出装置5は、式(11)を用いてクラス帰属確率を算出する(ステップA6)。このクラス帰属確率は、 j 番目の部分空間に含まれるピクセルが、 i 番目のクラスに属する確率を表す。したがって、画像各ピクセルに対してこのクラス帰属確率を算出し、確率の高いクラスに各ピクセルを分類することによって、領域抽出が容易に行うことが可能となる。クラス帰属確率算出装置5は、算出したクラス帰属確率をパラメータ更新装置6に送る。

【0135】

パラメータ更新装置6は、式(9)を最大化するようにパラメータを更新する。具体的には、以下のようにパラメータを更新する。

【0136】

【数 19】

$$w_i = \sum_{j=1}^N \tilde{q}_j \tilde{\pi}_{ij} \quad \dots\dots(16)$$

$$\mu_i = \frac{\sum_{j=1}^N \tilde{q}_j \tilde{\pi}_{ij} \tilde{x}_j}{\sum_{j=1}^N \tilde{q}_j \tilde{\pi}_{ij}} \quad \dots\dots(17)$$

$$\Sigma_i = \frac{\sum_{j=1}^N \tilde{q}_j \tilde{\pi}_{ij} (\tilde{x}_j - \mu_i) \otimes (\tilde{x}_j - \mu_i)^T}{\sum_{j=1}^N \tilde{q}_j \tilde{\pi}_{ij}} \quad \dots\dots(18)$$

ここで、

【0137】

【数 20】

$$u \otimes v^T$$

は、ベクトル u , v の i 成分と j 成分の積 $u_i v_j$ を ij 成分に持つマトリクスを表す。
また、

【0138】

【数 21】

$$\tilde{x}_j$$

は、式 (14) で定義される、 j 番目の部分空間 A_j に含まれるデータの平均値である。

【0139】

画像がグレースケールの場合、 w_i 、 μ_i の更新は、 x として輝度値を用いて同様に行うことができ、標準偏差 σ の更新は、

【数 2 2】

$$\sigma_i^2 = \frac{\sum_{j=1}^N \tilde{q}_j \tilde{\pi}_{ij} (\tilde{x}_j - \mu_i)^2}{\sum_{j=1}^N \tilde{q}_j \tilde{\pi}_{ij}}$$

となる。

すでに述べたように、このようにパラメータを更新すれば、式(9)で与えられる粗視化された完全対数尤度は増加し、したがって式(12)で与えられる粗視化平均対数尤度も増加する(非特許文献3)。パラメータ更新装置6は、パラメータを更新した後、更新したパラメータを評価関数算出装置7に送る。評価関数算出装置7は、式(12)を用いて粗視化対数尤度を算出する(ステップA8)。

【0140】

また、評価関数としては、式(12)で示した粗視化対数尤度の他に、以下に説明する赤池の情報量基準(AIC: Akaike Information Criterion)、MDL、またはストラクチャル・リスクを用いることもできる。

まず、AICを説明する。

【0141】

【数 2 3】

$$L_{AIC} = -2n\tilde{L} + 2m \quad \cdots \cdots (19)$$

ここで、AICが小さいほど良い推定結果であるとする評価関数を用いることも可能である。ここでmは、全パラメータの個数を表す。AICは粗視化対数尤度にマイナスを掛けた量に比例するため、パラメータ変更装置6で行う更新によってAICは減少する方向に変化する。また、パラメータに比例する項が付加されているため、同じ粗視化対数尤度ならばパラメータの少ないモデルを用いた推定結果の方が良いとする。この評価関数を用いることにより、データへの過剰な適合が抑制され、雑音に強い推定を行うことが可能となる。

【0142】

また、以下のMDL、

【0143】

【数 2 4】

$$L_{MDL} = -n\tilde{L} + \frac{m}{2} \log n \quad \cdots \cdots (20)$$

を用いても同様な効果が得られる。

【0144】

さらに、ストラクチャル・リスク (Structural Risk)

【0145】

【数 25】

$$L_{SRM} = \frac{\bar{L}}{1 - c \sqrt{a_1 \frac{h(\ln a_2 n / h + 1) - \ln(\eta / 4)}{n}}} \quad \cdots \cdots (21)$$

を用いても同様な効果が得られる。ここで η は式 (21) が確率 η で成り立つことを表し、通常 0.01 等の値を用いる。c, a_1 , a_2 は、確率分布の性質によって決まる定数であるが、通常 $c=1$, $a_1=1$, $a_2=1$ 等の値が用いられる。h は VC 次元と呼ばれるものであり、パラメータの数に比例する量である。

【0146】

評価関数算出装置 7 は、評価関数の変化が所与の終了条件を満たしているか否かを調べ、終了条件が満たされていれば領域抽出装置 8 に現在のパラメータを送り、終了条件が満たされていなければ粗視化条件付き確率分布算出装置 4 に現在のパラメータを送る (ステップ A9)。終了条件としては、たとえば現時点での評価関数値と前回算出した評価関数値の差を現時点での評価関数値で割り、その値の絶対値が 0.0001 以下であるか否か、等の条件を用いることができる。

【0147】

領域抽出装置 8 は、評価関数算出装置 7 からパラメータを受け取り、パラメータの情報を用いて領域を抽出する (ステップ A10)。たとえば、i 番目のクラスに属する領域を抽出するには、クラス帰属確率の値を $j=1$ から $j=N$ まで調べ、その確率がある値 (閾値) 以上の部分空間を i 番目のクラスに属する部分空間とする。次に、その部分空間に含まれるピクセルを調べ、これらのピクセルを i 番目の領域であるとして抽出する。クラス帰属確率の閾値としては、たとえば 0.5 を用いれば、所望の結果を得ることができる。

【0148】

領域抽出装置 8 は、自動的に閾値を設定することも可能である。そのためには以下の処理を行う。i 番目の領域を抽出するためには、まず推定されたクラスの混合比率 w_i に全ピクセル数を掛け、各クラスに属するピクセル数の推定値を求める。この数を n_i とする。

【0149】

次に、クラス帰属確率を $j=1$ から $j=N$ まで調べ、値の高い部分空間から順に、その部分空間に含まれるピクセルを抽出し、抽出したピクセルが n_i に到達するまで続ける。 n_i 番目に抽出したピクセルの番号を l とすると、式 (11) の粗視化されたクラス帰属確率の値が、i 番目の領域に属する確率の閾値となる。このように自動的に設定された閾値を用いることにより、画像から染色体テリトリを精度良く抽出することができる。領域抽出装置 8 は、領域抽出が完了した後、抽出した領域のデータを出力装置 9 へ送る。

【0150】

以上で説明した、データ空間粗視化装置 2、粗視化経験分布算出装置 3、粗視化条件付き確率分布 4、クラス帰属確率算出装置 5、パラメータ更新装置 6、評価関数算出装置 7、領域抽出装置 8 は、たとえば、パーソナルコンピュータやワークステーション、あるいはスーパーコンピュータ等の計算機を用いて構築することができる。出力装置 9 は、領域抽出装置 8 から領域データを受け取り、図 2 の定量化処理部 106 に出力する (ステップ A11)。また、出力装置 9 からの出力は、出力部 112 によって表示することもできる。

【0151】

図 8 は、抽出処理部 104 の他の例を示すブロック図である。図 9 は、図 8 に示した抽出処理部 104 の処理手順を示すフローチャートである。以下、図 3 および図 4 を参照して説明したのと同様の構成要素には同様の符号を付し、適宜説明を省略する。

【0152】

ここで、領域復元装置 10 は、粗視化確率分布を用いたパラメータ推定が完了した後、

粗視化の解像度が元の解像度に等しいかを調べ（ステップA12）、粗視化されていれば、データを元の解像度に戻す（ステップA13）。粗視化されていなければ、元の解像度での推定が完了していることを意味するため、元の解像度で推定されたパラメータを領域抽出装置8に送る。データを元の解像度に戻すには、粗視化の解像度をデータ属性値の最小単位に設定し（たとえば1）、第1の実施の形態に述べた方法と全く同一の方法を繰り返せばよい。この場合、図3および図4を参照して説明した方法よりも推定に時間を要するようになるが、より高い精度でパラメータを推定することができ、結果としてより精度の高い領域抽出が可能となる。また、粗視化確率分布を用いて推定したパラメータは、すでに最適なパラメータの近傍に推定されているため、最初から高い解像度で推定を行うよりも少ない回数のパラメータ更新で最適パラメータが推定でき、はるかに高速に領域抽出を行うことができる。

【0153】

（配置状態定量化（S22））

図2に戻り、定量化処理部106は、染色体テリトリの配置状態の定量化に先立ち、抽出処理部104により抽出された染色体テリトリの座標を一定の基準に従って変換して標準化する処理を行う。このような処理を行うことにより、複数の検体について、統一した基準で、染色体テリトリの配置状態を定量化することができる。これにより、複数の検体における染色体テリトリの配置状態の定量化結果を統計処理等することができる。

【0154】

図10は、染色体テリトリの配置状態を定量化する処理手順を示すフローチャートである。図11は、染色体テリトリの座標を変換する様子を示す模式図である。以下、図10および図11を参照して座標変換の処理手順を説明する。

【0155】

定量化処理部106は、図11（a）に示すような細胞の核120を標準化する（図10、S100）。ここでは、まず、細胞の核120の重心 P_0 および体積 V を算出する。つづいて、図11（b）に示すように、重心 P_0 を中心とし、体積 V となる標準球124を算出する。ここで、細胞の核120には染色体テリトリ122が含まれる。ここでは図示していないが、細胞の核120は、複数の染色体テリトリを含むことができる。

【0156】

つづいて、細胞の核120および標準球124の形状およびサイズに応じて、染色体テリトリ122の座標を変換する（図10、S102）。具体的には、図11（c）に示すように、標準球124の中心（重心 P_0 と同じ）および染色体テリトリ122内の各点 p を通る直線 l を想定し、直線 l 上において、中心 P_0 と細胞の核120の表面までの距離 R を算出する。同時に、直線 l 上において、中心 P_0 と標準球124の表面までの距離 r_0 を算出する。ここで、図11（d）に示すように、染色体テリトリ122内の点 p を標準球120内の点に変換すると、標準球120内における座標 p' は、

$$p' = r_0 / R \times (p - P_0) + P_0$$

となる。染色体テリトリ122内のすべての点に対して同様に変換することにより、標準球124内に座標変換された染色体テリトリ126を得ることができる。

【0157】

本実施の形態において、定量化処理部106は、以上のような染色体テリトリの座標変換を行った後、染色体テリトリの配置状態を定量化する（図10、S104）。

【0158】

図12は、標準球124内に座標変換された複数の染色体テリトリ126aおよび染色体テリトリ126bを示す図である。ここでは、配置状態の定量化結果として、複数の染色体テリトリ間の距離を用いる例を示す。定量化処理部106は、このように座標変換された染色体テリトリ126aおよび染色体テリトリ126b間の距離を算出する。

【0159】

以上のような染色体テリトリの座標変換処理を行うことにより、種々の形状の複数の検体における染色体テリトリの配置状態の定量化結果を統一された指標で処理することがで

きる。また、細胞の核は、同じ細胞であっても、時期によって種々の形状をとり得る。本実施の形態において、このような座標変換処理を行うことにより、このような時期の違いによる細胞の核の形状の変化の影響を排除して、染色体テリトリの配置状態を適切に定量化することができる。

【0160】

さらに、配置状態の定量化結果として染色体テリトリ間の距離を用いる場合、定量化処理部106は、染色体テリトリ間の距離を標準球124の半径で除した値を定量化の結果とすることができる。これにより、それぞれ核の大きさが異なる複数の検体における染色体テリトリの配置状態の定量化結果を統一された指標で処理することができる。

【0161】

次に、染色体テリトリの配置状態を定量化する方法を説明する。

染色体テリトリの配置状態は、複数の染色体テリトリ間の距離、染色体テリトリの広がり、配置方向、および染色体テリトリと核内の基準点との距離等を用いて定量化することができる。

【0162】

複数の染色体テリトリ間の距離を用いて定量化を行う場合、複数の染色体テリトリ間の距離は、種々の方法で算出したものを用いることができる。

【0163】

たとえば、定量化処理部106は、各染色体テリトリの重心をそれぞれ算出し、重心間の距離を算出することができる。ここで、重心とは、染色体テリトリを構成する各ピクセルの比重が等しいと仮定した面積の重心のことである。

各染色体テリトリの重心位置としては、例えば輝度の値で重みをとった重心

$$X_G = \sum_j I_j x_j / \sum_j I_j$$

を用いることもできる。ここで、 I_j は位置 x_j におけるピクセルの輝度値である。また、重心の代わりに最も輝度値が高いピクセルの位置を用いてもよい。以下では、単純な重心を例に説明するが、重み付き重心や最大輝度のピクセル位置を用いた場合でも同様な方法で染色体間の距離を定義することができる。

【0164】

また、定量化処理部106は、複数の染色体テリトリの境界間の距離を算出することもできる。ここで、染色体テリトリの境界間の距離とは、たとえば、複数の染色体テリトリの表面の最短距離である。

【0165】

また、定量化処理部106は、複数の染色体テリトリの形状をそれぞれ正規分布で近似し、これらの染色体テリトリそれぞれの正規分布の平均値を算出し、平均値間の距離を算出することもできる。また、複数の染色体テリトリの形状をそれぞれ正規分布で近似した後に上述したように各染色体テリトリの重心を算出して重心間の距離を算出したり、境界間の距離を算出することもできる。

【0166】

また、定量化処理部106は、複数の染色体テリトリそれぞれの骨格線を検出し、各骨格線の重心を算出し、重心間の距離を算出することもできる。

【0167】

また、定量化処理部106は、複数の染色体テリトリ間の距離を、マハラノビス (Mahalanobis) の距離、バタチャリア (Bhattacharyya) の距離、カルバック-ライブラー (Kullback-Leibler) の距離、対称カルバック-ライブラーの距離、および絶対距離により算出することもできる。

【0168】

マハラノビスの距離は、以下の式に従って算出することができる。

【0169】

【数 2 6】

$$D = (\mu_1 - \mu_2)^T \Sigma^{-1} (\mu_1 - \mu_2)$$

ここで、 μ_i ($i = 1, 2$) は平均ベクトル、 Σ は、

【0 1 7 0】

【数 2 7】

$$\Sigma = w_1 \Sigma_1 + w_2 \Sigma_2$$

で定義されるクラス内共分散行列、 w_1 、 w_2 は、クラスの混合比率を示す。

【0 1 7 1】

バタチャリアの距離は、以下の式に従って算出することができる。

【0 1 7 2】

【数 2 8】

$$D = \frac{1}{8} (\mu_1 - \mu_2)^T \left[\frac{\Sigma_1 + \Sigma_2}{2} \right]^{-1} (\mu_1 - \mu_2) + \frac{1}{2} \ln \frac{|\frac{\Sigma_1 + \Sigma_2}{2}|}{\sqrt{|\Sigma_1| |\Sigma_2|}}$$

【0 1 7 3】

カルバック-ライブラーの距離は、以下の式に従って算出することができる。

【0 1 7 4】

【数 2 9】

$$D = \int p_1(x) \ln \left[\frac{p_1(x)}{p_2(x)} \right] dx$$

ここで、 p_i ($i = 1, 2$) は、平均ベクトル μ_i ，共分散行列 Σ_i の多次元正規分布を示す。

【0 1 7 5】

対称カルバック-ライブラーの距離は、以下の式に従って算出することができる。

【0 1 7 6】

【数 3 0】

$$D = \frac{1}{2} \int p_1(x) \ln \left[\frac{p_1(x)}{p_2(x)} \right] dx + \frac{1}{2} \int p_2(x) \ln \left[\frac{p_2(x)}{p_1(x)} \right] dx$$

【0 1 7 7】

絶対距離は、以下の式に従って算出することができる。

【0 1 7 8】

【数 31】

$$D = \int |p_1(x) - p_2(x)| dx$$

【0179】

染色体テリトリの広がりを用いて定量化を行う場合、たとえば、染色体テリトリの形状を正規分布を用いて近似する方法を用いることができる。これにより、染色体テリトリの広がりの方角と大きさを定量化することができる。具体的には、もとの画像に対し、輝度の値を染色体テリトリの存在する確率の大きさと見なし、混合正規分布で近似する。この処理は、上述した抽出処理部 104 による染色体テリトリの抽出方法と同様の手順で行うことができる。この際、画像上の各ピクセルの座標を 2 次元の確率ベクトルと見なし、その位置における輝度の値を確率値と見なす。なお、輝度を確率値として扱うために、画像中の全ピクセルの輝度値の総和を求め、その値で各ピクセルの輝度値を除し、正規化しておく。このようにして、染色体テリトリの形状を定量化する。

【0180】

染色体テリトリは、細胞の核内においてある程度広がりをもって存在しており、この広がり形状は染色体どうし、あるいは核膜を通した外部との相互作用の生じ易さと関係があると考えられる。従って、染色体テリトリの配置状態を定量化する際に、染色体テリトリの広がり形状を考慮することにより、細胞の状態を精度良く評価できることが期待される。

【0181】

染色体テリトリ間の配置方向を用いて定量化を行う場合、各染色体テリトリの主軸を検出し、それらの主軸がなす角度を定量化した結果とすることができる。染色体テリトリの主軸も、たとえば上述したような正規分布を用いて算出することができる。このように、染色体テリトリの形状を正規分布で近似することにより、各染色体テリトリの広がり主軸を明瞭にすることができる。

【0182】

このとき、推定されるパラメータは、

$$P(w_1, \mu_1, \Sigma_1, w_2, \mu_2, \Sigma_2) = w_1 \times \Phi(\mu_1, \Sigma_1) + w_2 \times \Phi(\mu_2, \Sigma_2)$$

となる。ここで $\Phi(\mu, \Sigma)$ は 2 次元正規分布であり、 μ は平均値、 Σ は共分散行列である。この結果得られたパラメータ w は広がり大きさの比率、 μ は広がり中心位置、 Σ は広がり主軸を表す。さらに、共分散行列を固有値分解し、広がり主軸を具体的に得る。以上のように、形状を定量化することにより、染色体テリトリどうしの配置状態を数値化することができる。

【0183】

二つの染色体テリトリの配置状態を考慮するとき、二つの染色体テリトリの広がり主軸が平行である場合の方が、垂直である場合よりも、互いに相互作用する面積が大きく、影響を及ぼしやすいと考えられる。従って、染色体テリトリの配置状態を定量化する際には、染色体テリトリの広がり形状や配置方向を考慮することにより、細胞の状態をより精度良く評価できることが期待される。

【0184】

染色体テリトリと核内の基準点との距離を用いて定量化を行う場合、核内の基準点は、核膜や重心とすることができる。また、この場合、核膜とは、上述した方法で算出した標準球の表面とすることができる。一般に、染色体が細胞の核内の外縁部に存在している方が中心部に存在しているときよりも遺伝的に活性な状態にあることが知られている。従っ

て、染色体テリトリと細胞の核膜との距離や染色体テリトリと細胞の核膜との距離を定量化することにより、その細胞の遺伝的な活性を評価できることが期待される。

【0185】

定量化処理部106は、以上の方法を単独または複数用いて染色体テリトリの配置状態を定量化することができる。複数の方法で定量化した結果を用いることにより、細胞の状態をより詳細に評価することができる。

【0186】

(評価(S26))

細胞の状態の評価は、種々の方法で行うことができるが、複数の検体について得られた定量化結果を統計処理しておき、統計処理を行った結果と比較することにより、ある検体の細胞の状態を評価することができる。

【0187】

統計処理は、種々の方法を用いることができるが、たとえば統計処理部108は、複数の検体の染色体テリトリの配置状態を定量化した結果の標準値を算出することができる。この場合、標準値としては、配置状態を定量化した結果の平均値を用いることができる。このようにして標準値を算出した場合、ある検体の染色体テリトリの配置状態を定量化した結果を標準値と比較することにより、その検体の細胞の状態を評価することができる。この場合、評価装置100は、図13に示すように、図2に示した構成に加えて、統計処理部108および統計結果記憶部116を含むことができる。統計処理部108は、たとえば上記の標準値を算出し、統計結果記憶部116に標準値を保存する。評価部110は、ある検体の定量化結果を標準値と比較することにより、その検体の細胞の状態を評価することができる。

【0188】

また、統計処理は、各検体の染色体テリトリの配置状態を定量化した結果と、その検体の提供者の実際の診断結果とを対応付けて記憶する処理とすることもできる。たとえば、癌等の疾患の罹患者から採取した細胞の染色体テリトリの配置状態の定量化データと、非罹患者から採取した細胞の染色体テリトリの配置状態の定量化データとを準備しておくことができる。評価部110は、ある被験者から採取した細胞の染色体テリトリの配置状態の定量化結果が、いずれの定量化データに近いかを判断することにより、その被験者が罹患しているか否かを評価することができる。

【0189】

本実施の形態において、染色体テリトリの配置状態を定量化した結果を用いて細胞の状態を評価し、その結果に基づいて疾患の有無を判断する。染色体テリトリの配置状態の変化は、細胞に形態的变化が生じるよりも早い段階で生じるため、疾患の早期発見が期待できる。

【0190】

また、統計処理部108は、複数の検体から得られた定量化結果に基づき、ニューラルネットワークを構築することができる。ここで、たとえば、複数の染色体テリトリ間の距離、染色体テリトリの広がり、配置方向、および染色体テリトリと核内の基準点との距離等、種々の方法で定量化した複数の結果を入力パラメータとすることができる。評価部110は、ニューラルネットワークの出力により細胞の状態を評価することができる。

【0191】

さらに、評価部110は、t検定やカイ2乗検定を用いて一の検体の配置状態を定量化した結果と統計処理の結果との差異を定量化することができる。

【実施例】

【0192】

脂肪細胞(25番細胞および13番細胞)を用い、3次元のカラー画像(554×576)を得た。ここで、カラー画像は、それぞれ、16番染色体、12番染色体、および核と相補的に結合するプローブを用いて3D-FISH法により得た。

【0193】

画像取得部102により、このようなカラー画像を入力した。抽出処理部104は、粗視化解像度を8とし、クラスの数として上述した粗視化領域抽出法を用いて画像から染色体テリトリおよび核を抽出した。クラスは、12番染色体の染色体テリトリ、16番染色体の染色体テリトリ、核それぞれに割り当てた。

【0194】

この結果、図14に示すように、12番染色体の染色体テリトリ（クラス1）、16番染色体の染色体テリトリ（クラス2）、細胞核（クラス3）が抽出された。図14は、抽出した画像を示す。

【0195】

つづいて、定量化処理部106により、16番染色体の染色体テリトリおよび12番染色体の染色体テリトリの重心間の距離を算出した。16番染色体および12番染色体には、相同染色体が含まれるが、それらのうち、16番染色体および12番染色体間の最短距離を採用した。染色体テリトリの距離が最も近い箇所で転座等の相互作用が生じると考えられるからである。

【0196】

まず、画像データから抽出された細胞核の標準球を算出し、12番染色体の染色体テリトリおよび16番染色体の染色体テリトリを規定化した。ここで、染色体テリトリは、ほぼ平坦な形状を有していたため、z方向における変換は行わず、二次元の変換を行った。三次元の変換を行う場合も同様の処理手順とすることができる。

【0197】

このようにして変換を行った後、25番細胞および13番細胞における、12番染色体と16番染色体の重心間の距離をそれぞれ算出した。その結果、25番細胞における重心間の距離は3.38 μm 、13番細胞における重心間の距離は4.12 μm であった。

【0198】

以上、本発明を実施の形態および実施例に基づいて説明した。この実施の形態および実施例はあくまで例示であり、種々の変形例が可能なこと、またそうした変形例も本発明の範囲にあることは当業者に理解されるところである。

【図面の簡単な説明】

【0199】

【図1】本発明の実施の形態における細胞の状態の評価方法の処理手順を示すフローチャートである。

【図2】図1に示した処理手順を実行する本実施の形態における評価装置を示すブロック図である。

【図3】図2に示した抽出処理部を示すブロック図である。

【図4】抽出処理部の処理手順を示すフローチャートである。

【図5】粗視化したときの粗視化データ空間および部分空間を示す図である。

【図6】粗視化経験分布に基づいて、パラメータの初期値を決定する処理手順を示したフローチャートである。

【図7】図6を参照して説明した手順を具体的に説明する図である。

【図8】抽出処理部の他の例を示すブロック図である。

【図9】図8に示した抽出処理部の処理手順を示すフローチャートである。

【図10】染色体テリトリの配置状態を定量化する処理手順を示すフローチャートである。

【図11】染色体テリトリの座標を変換する様子を示す模式図である。

【図12】標準球内に座標変換された複数の染色体テリトリおよび染色体テリトリを示す図である。

【図13】評価装置の他の例を示すブロック図である。

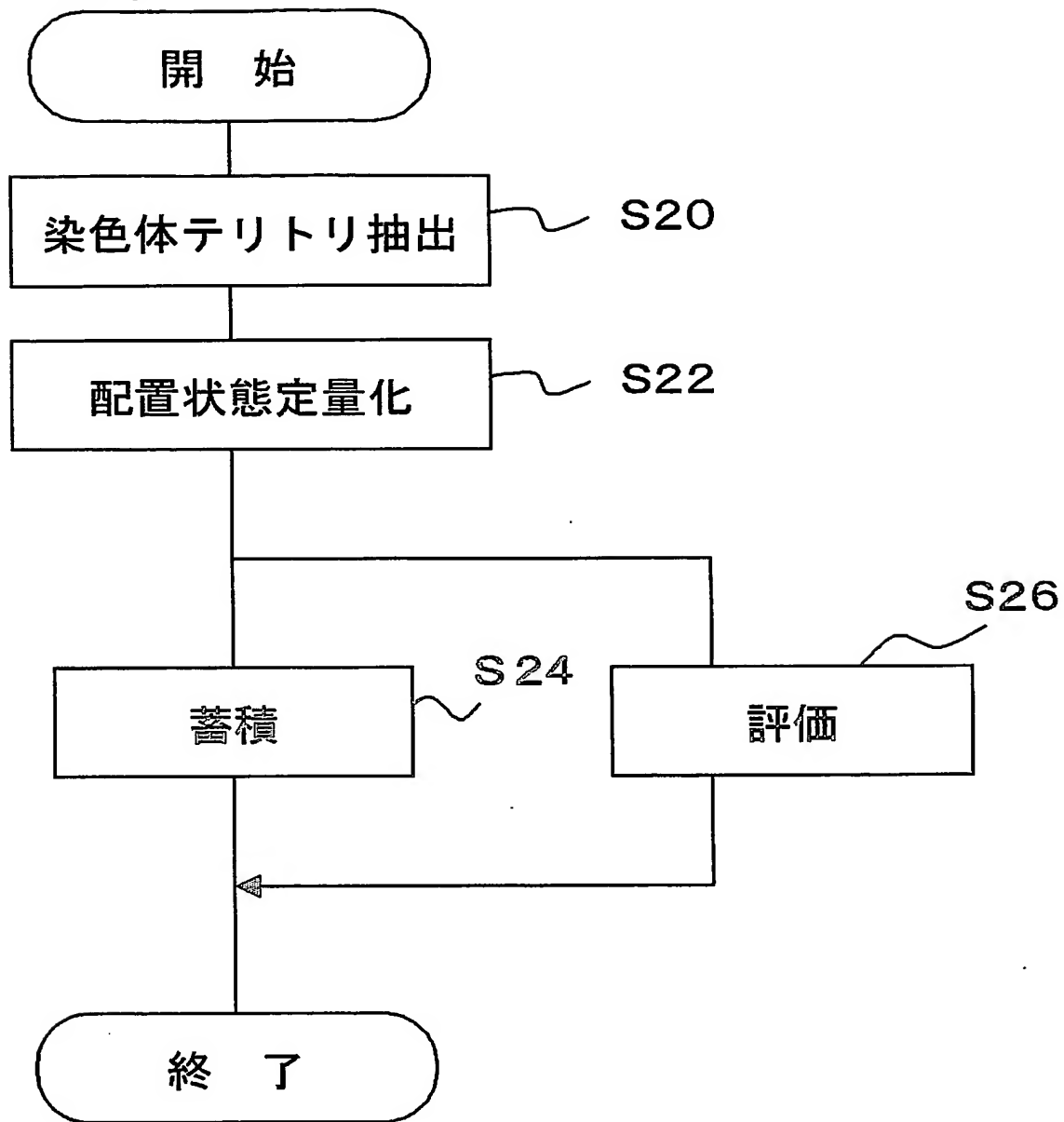
【図14】本実施の形態における抽出処理部により抽出された染色体テリトリを示す図である。

【符号の説明】

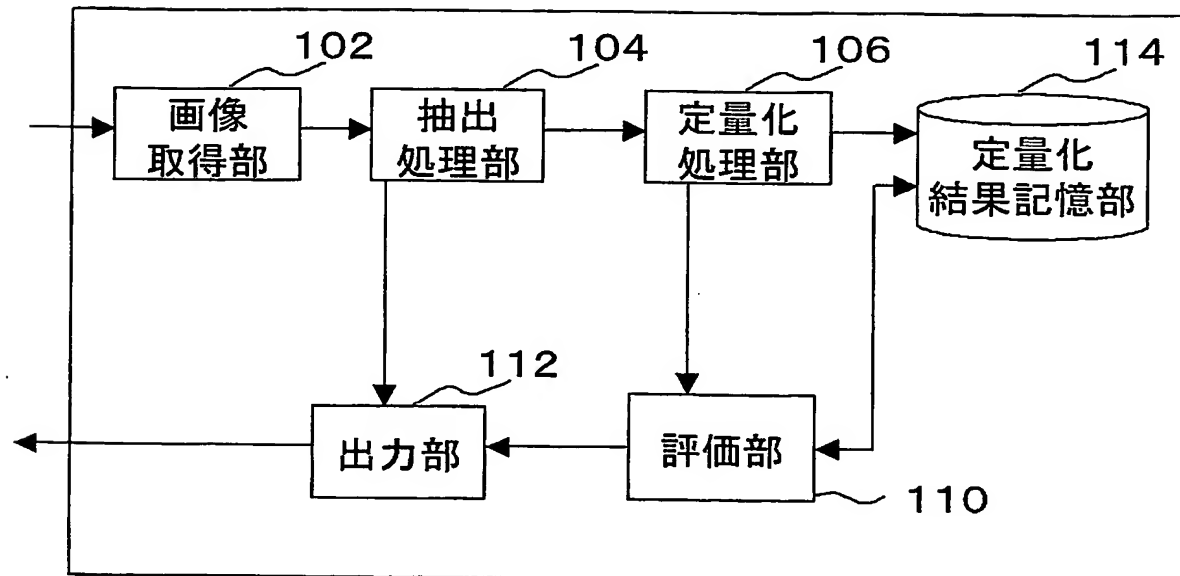
【 0 2 0 0 】

- 1 入力装置
- 2 データ空間粗視化装置
- 3 粗視化経験確率分布算出装置
- 4 粗視化条件付き確率分布算出装置
- 5 クラス帰属確率算出装置
- 6 パラメータ更新装置
- 7 評価関数算出装置
- 8 領域抽出装置
- 9 出力装置
- 1 0 領域復元装置
- 1 1 粗視化データ空間
- 1 2 部分空間
- 1 0 0 評価装置
- 1 0 2 画像取得部
- 1 0 4 抽出処理部
- 1 0 6 定量化処理部
- 1 0 8 統計処理部
- 1 1 0 評価部
- 1 1 2 出力部
- 1 2 0 核
- 1 2 2 染色体テリトリ
- 1 2 4 標準球
- 1 2 6 染色体テリトリ (座標変換後)
- 1 2 6 a 染色体テリトリ (座標変換後)
- 1 2 6 b 染色体テリトリ (座標変換後)

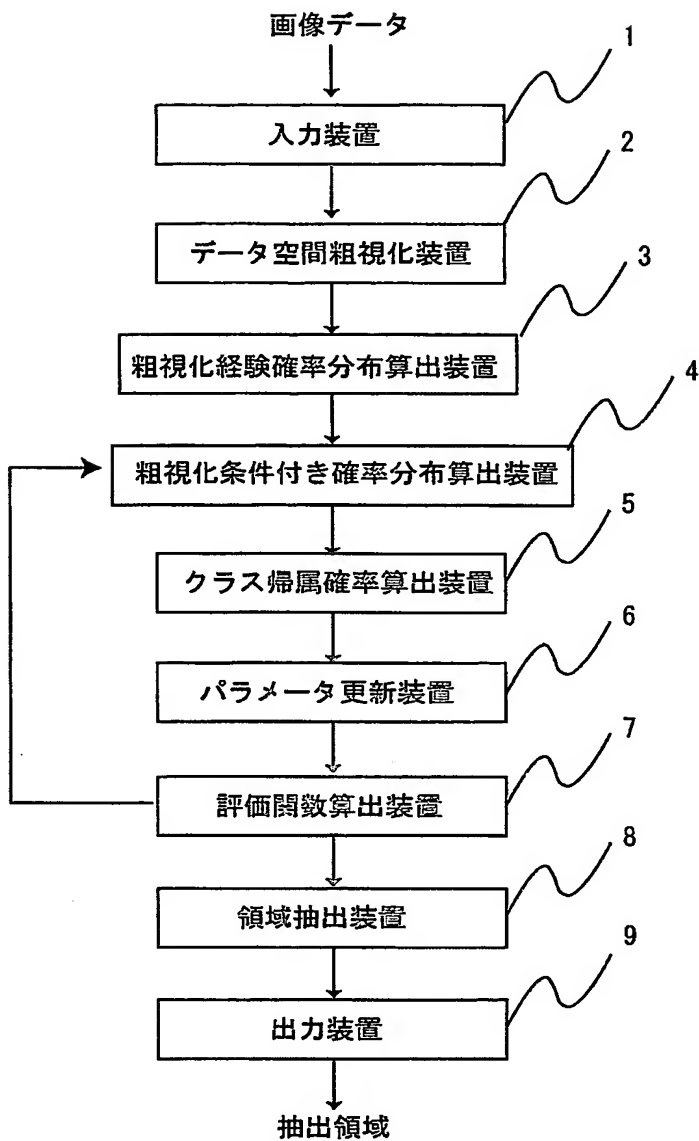
【書類名】 図面
【図 1】



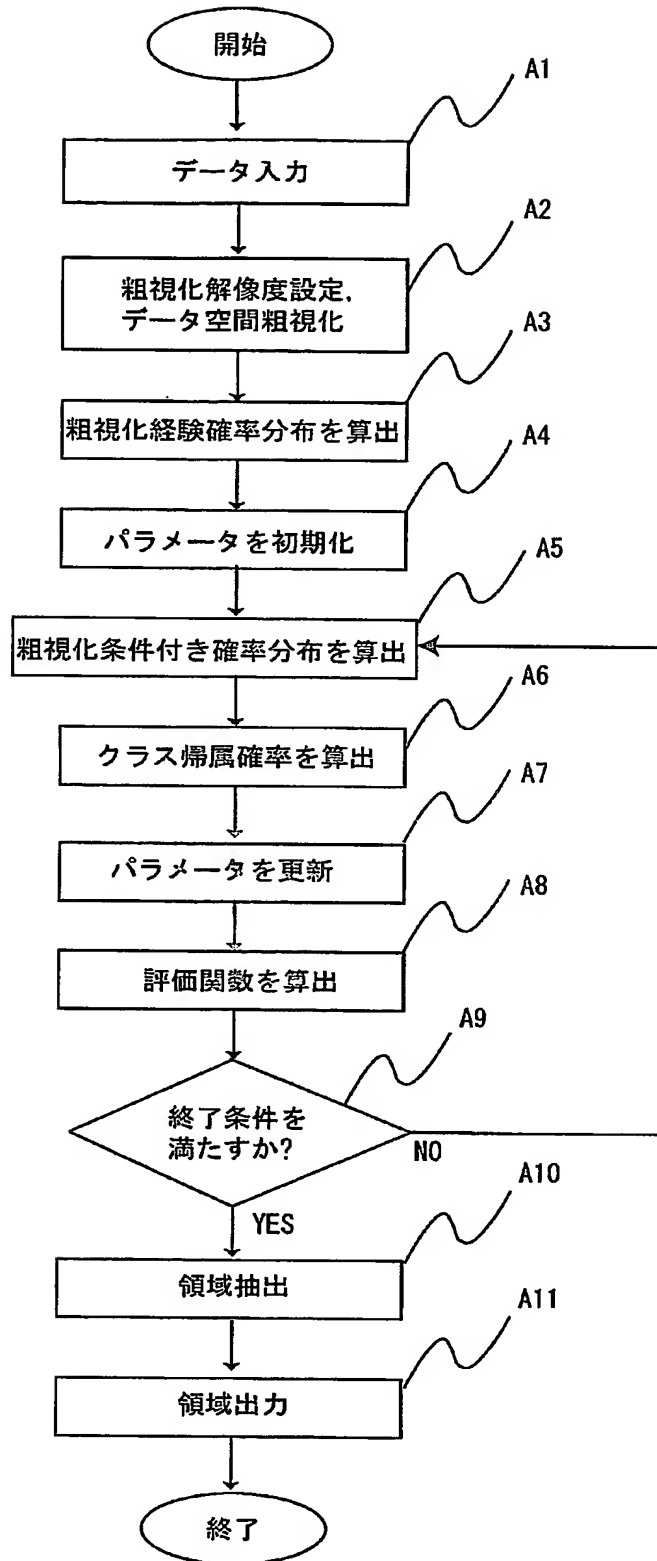
【図 2】



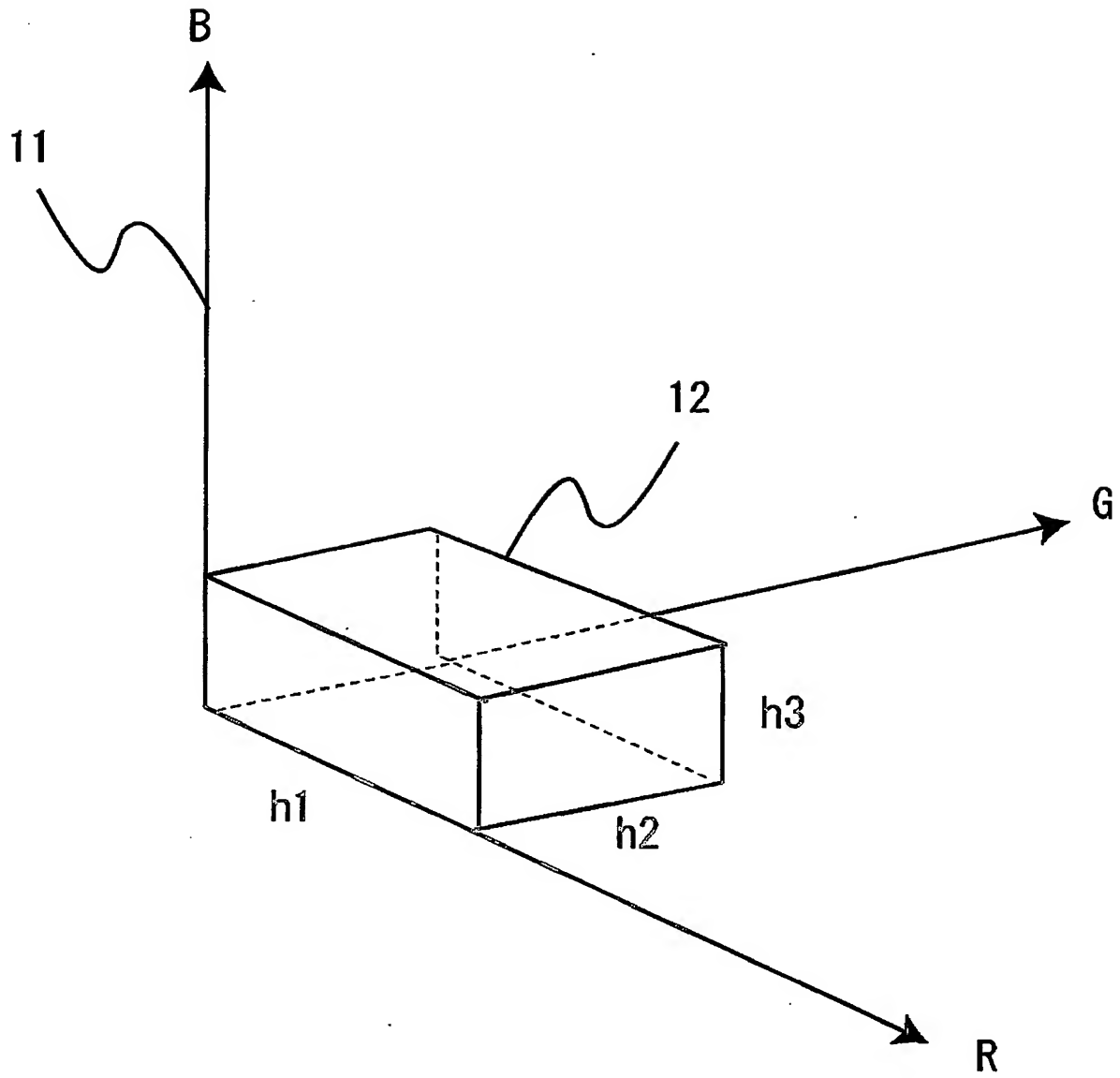
【図 3】



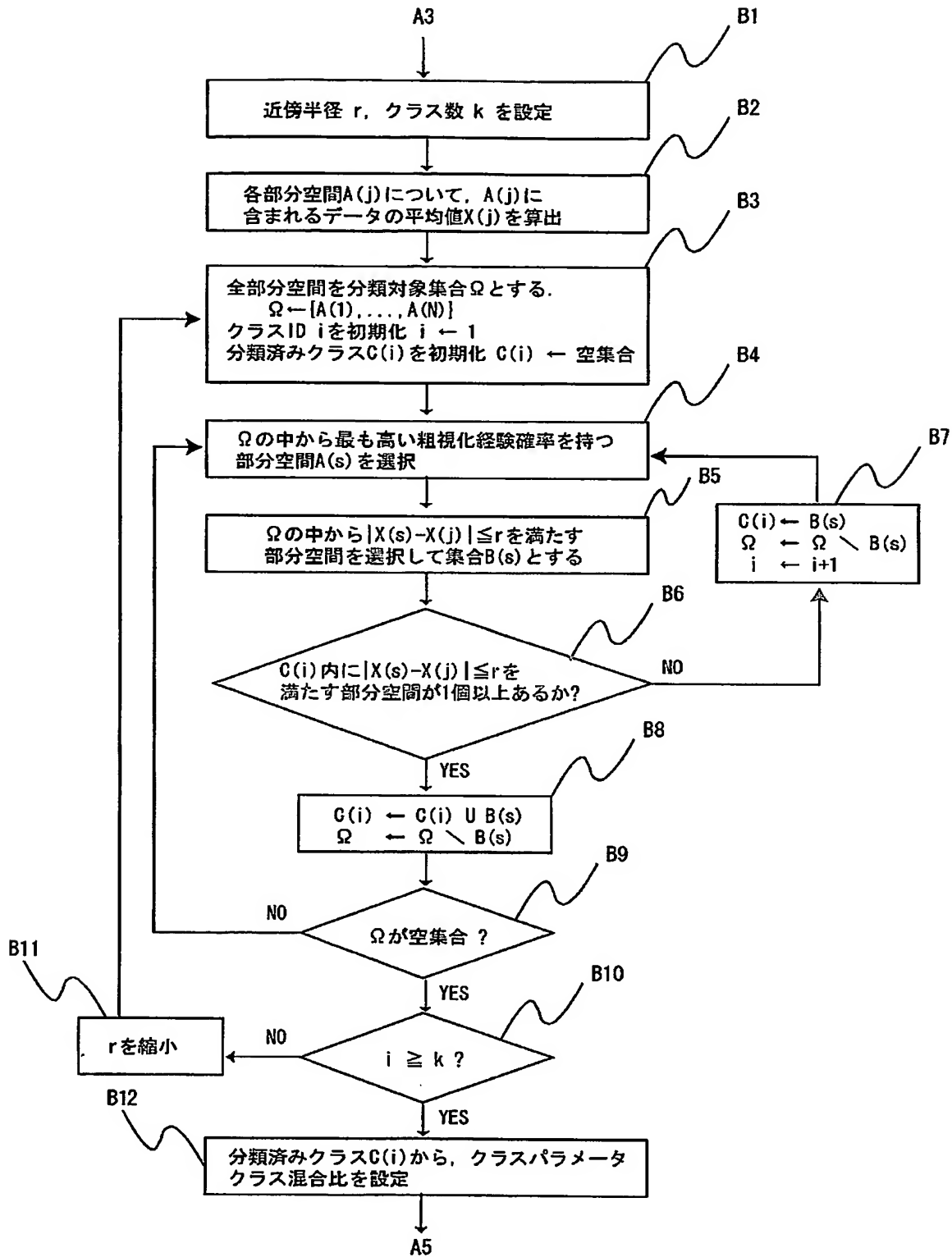
【図 4】



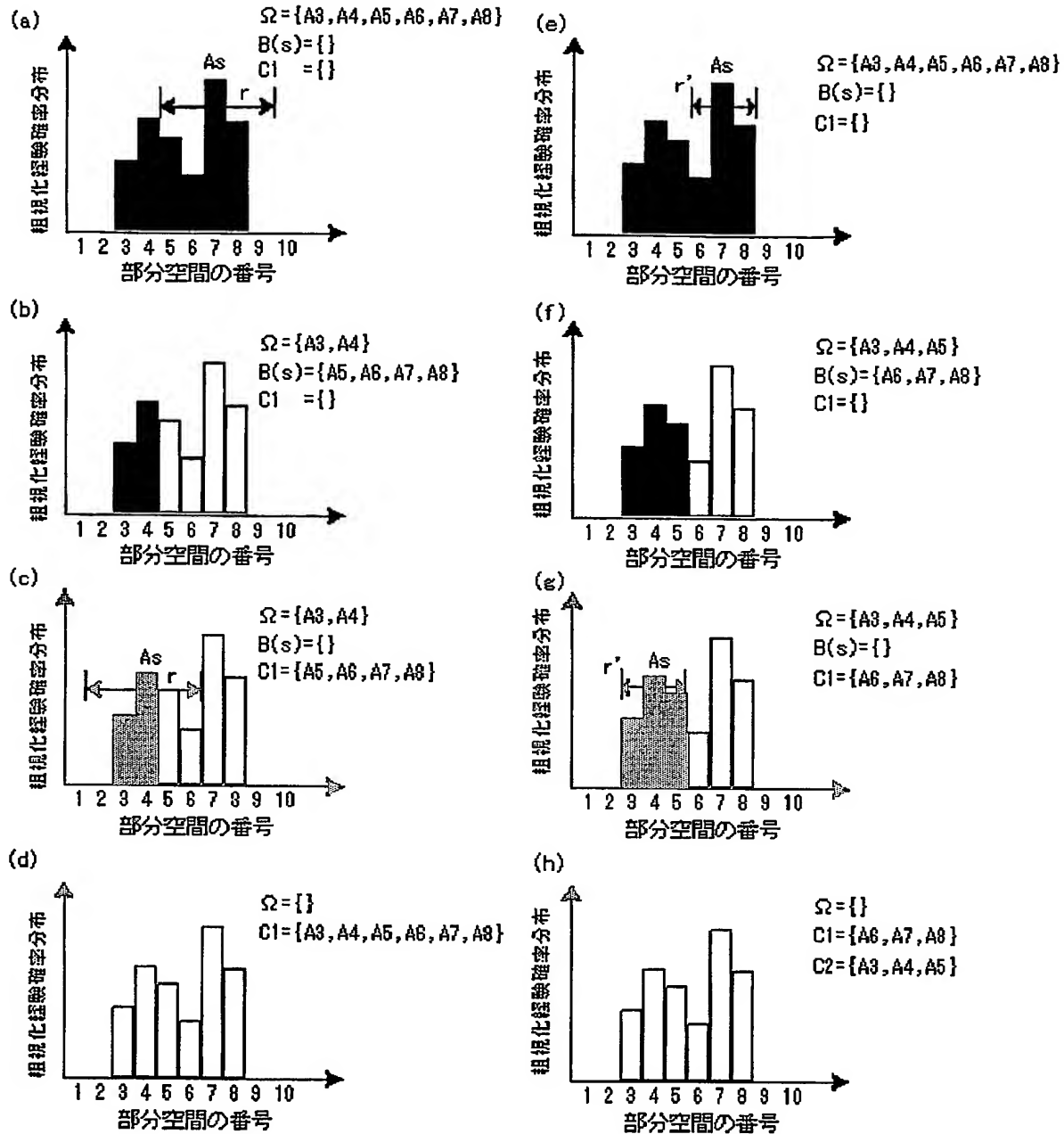
【図 5】



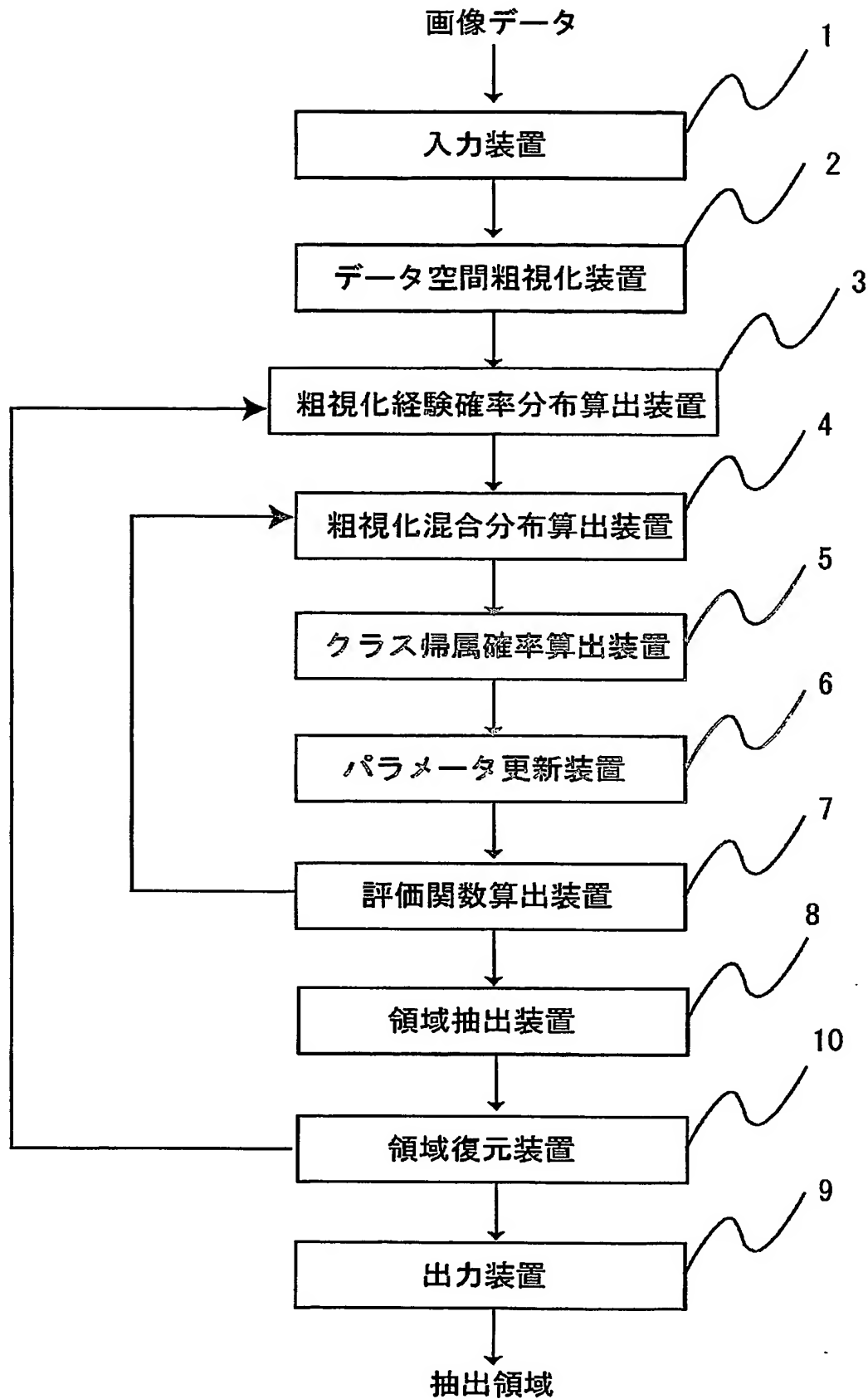
【図 6】



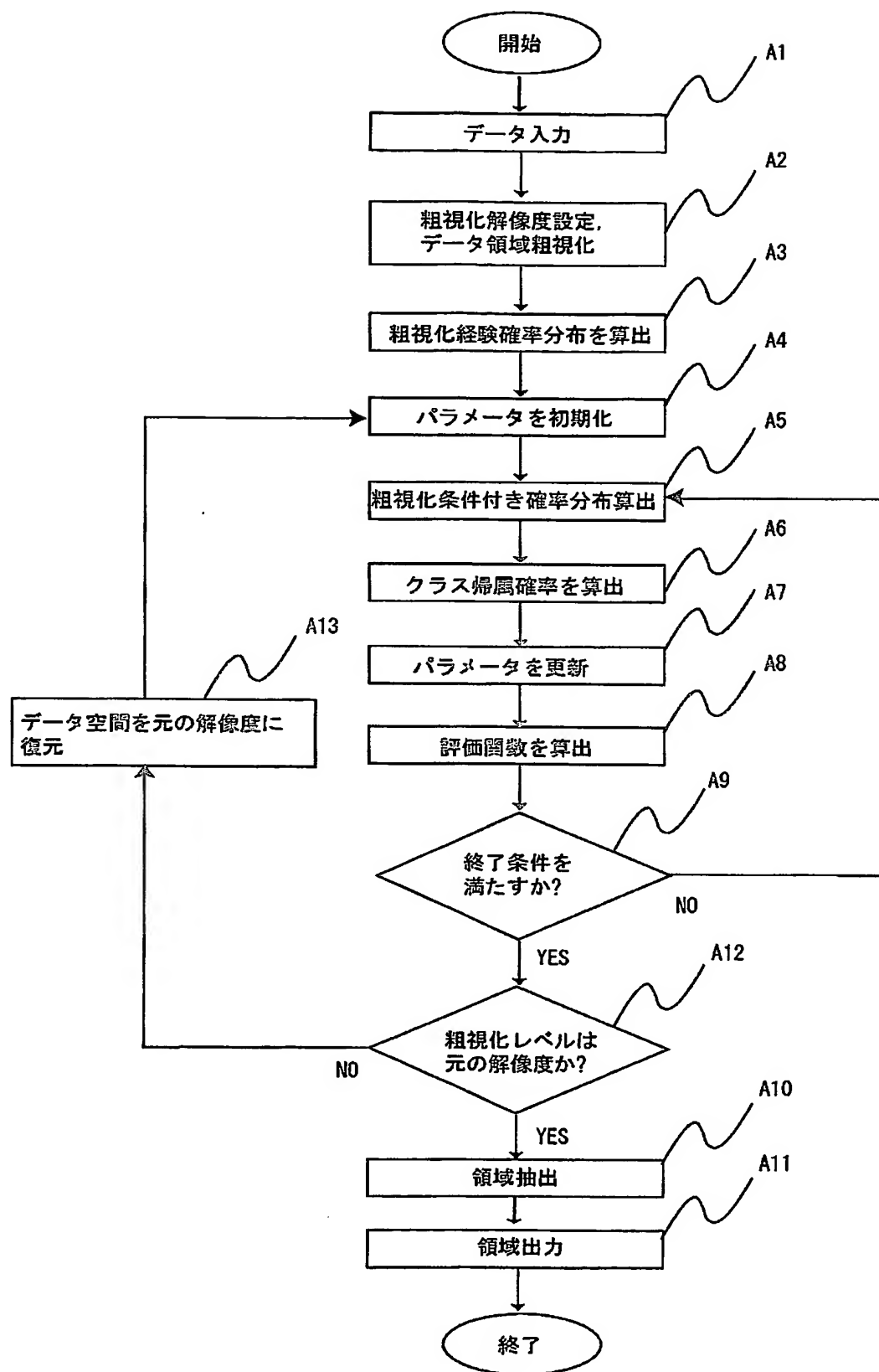
【図 7】



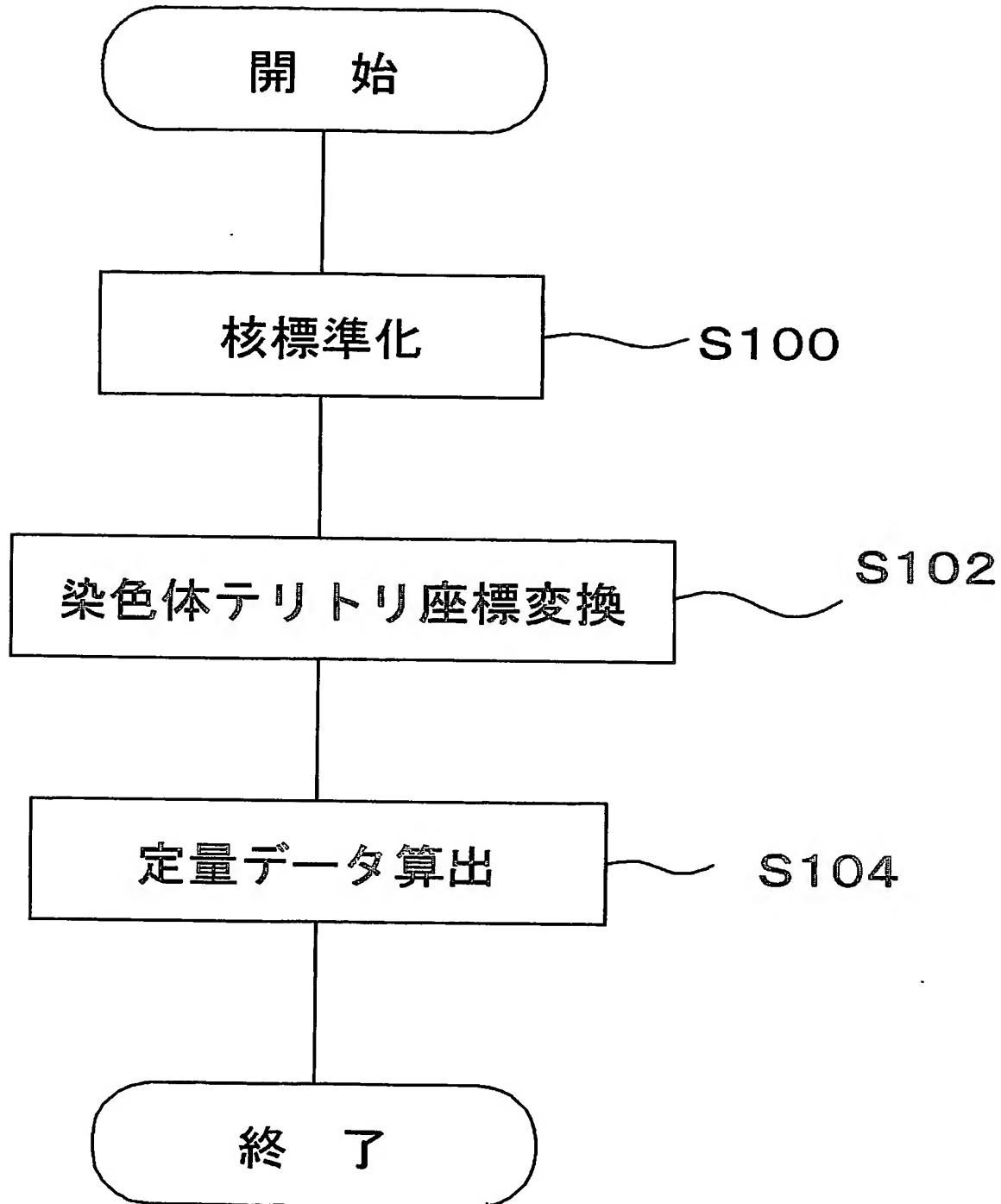
【図 8】



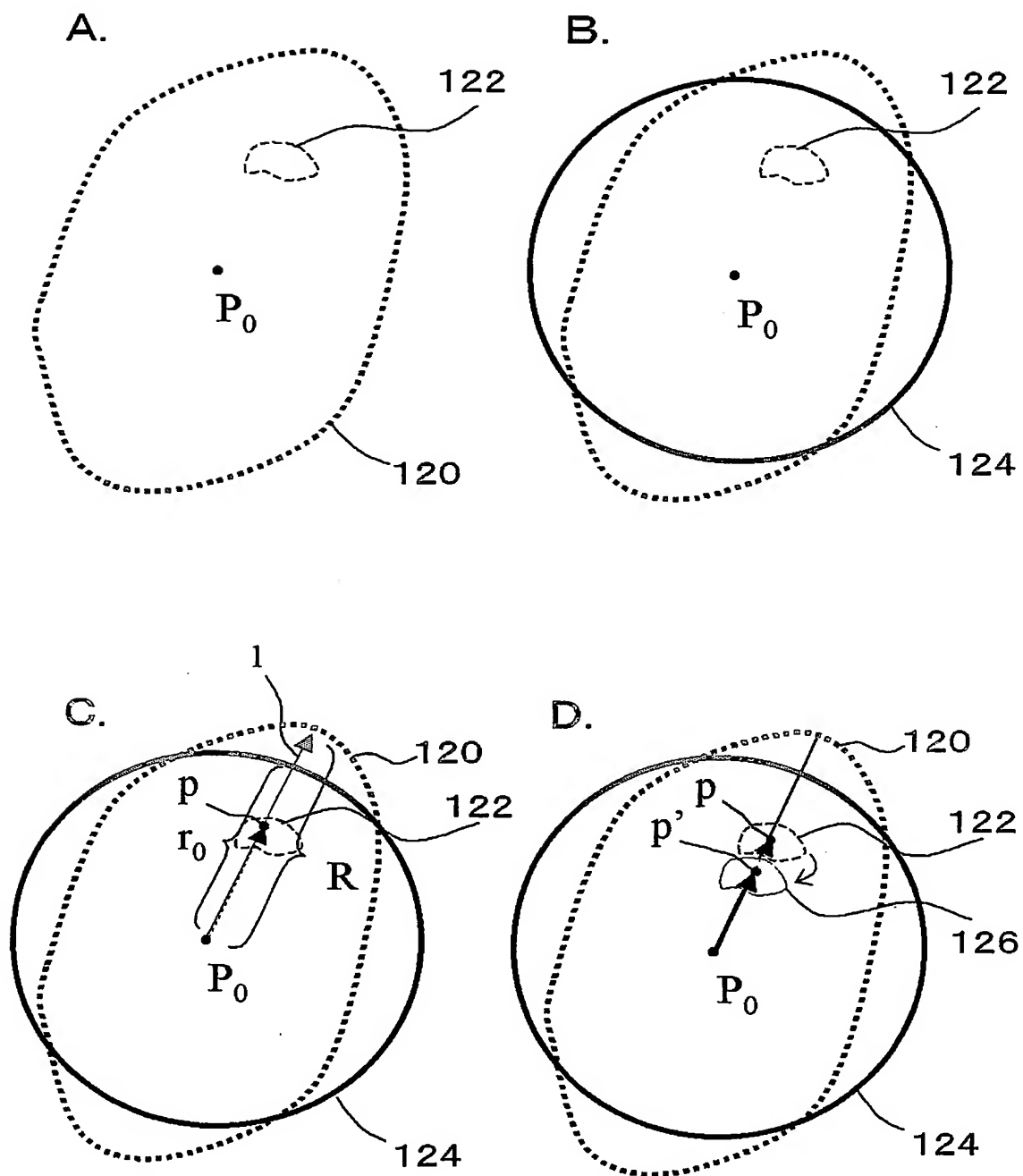
【図 9】



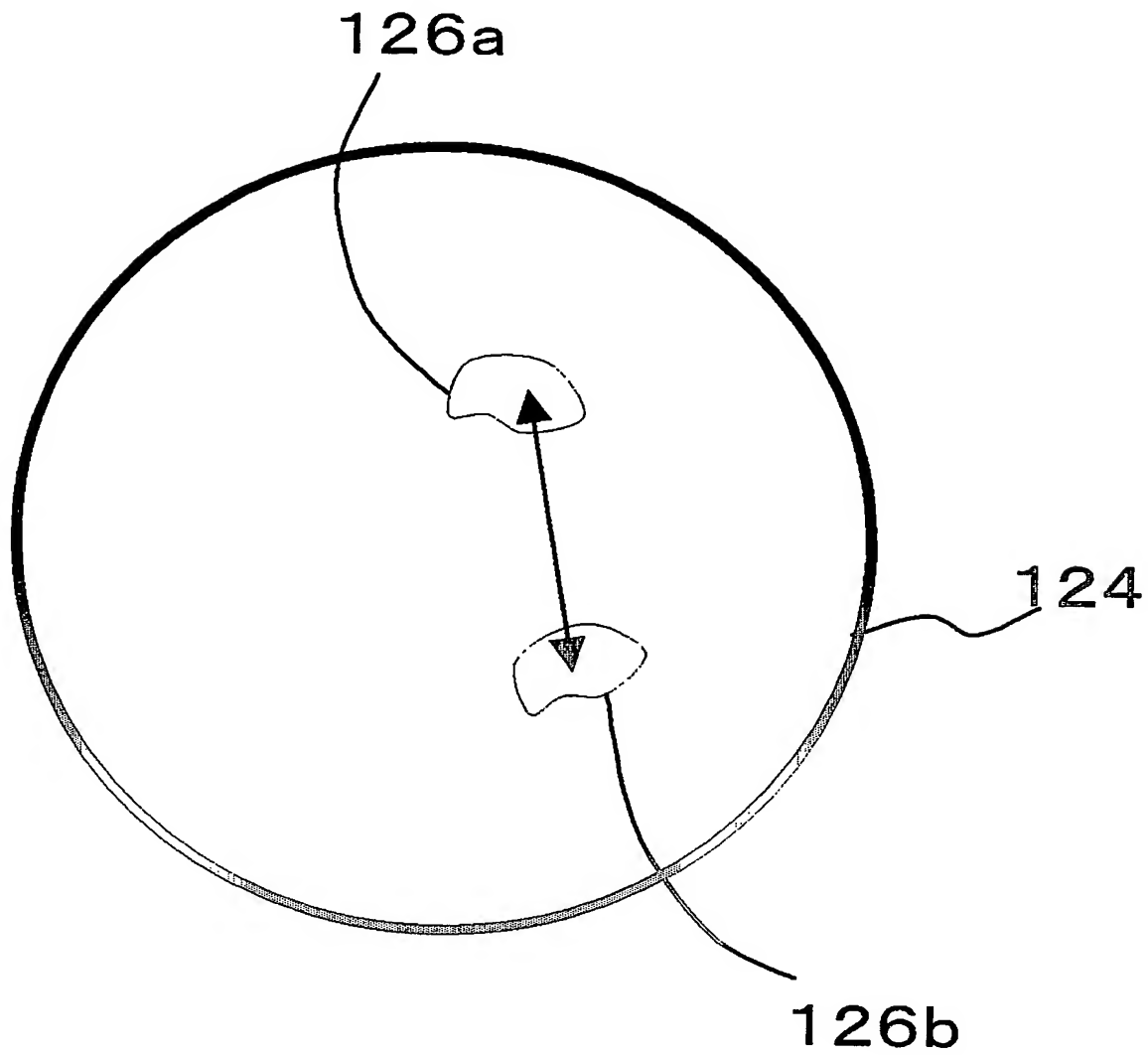
【図10】



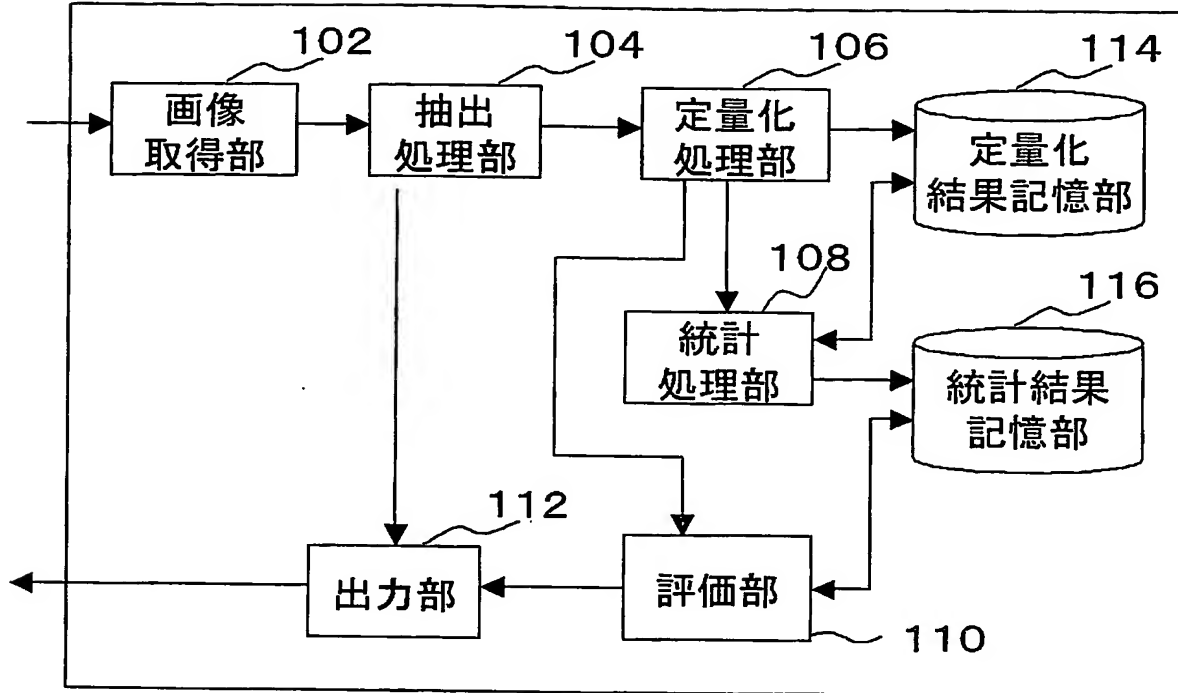
【図 11】



【図 12】



【図13】



【図 14】



BEST AVAILABLE COPY

【書類名】 要約書

【要約】

【課題】 細胞の状態を評価し、評価結果に基づき、癌等の疾患の診断の指標とする。

【解決手段】 染色体テリトリを含む細胞を撮影した画像情報に基づいて、細胞の状態を評価する方法を提供する。この方法は、前記画像から、前記染色体テリトリを抽出するステップ（S20）と、前記染色体テリトリの配置状態を標準化した後、当該配置状態を定量化するステップ（S22）と、定量化された前記染色体テリトリの配置状態に基づき、前記細胞の状態を評価するステップ（S26）と、を含む。

【選択図】 図1

特願 2 0 0 3 - 2 8 2 1 2 2

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 4 2 3 7]

1. 変更年月日

1 9 9 0 年 8 月 2 9 日

[変更理由]

新規登録

住 所

東京都港区芝五丁目 7 番 1 号

氏 名

日本電気株式会社